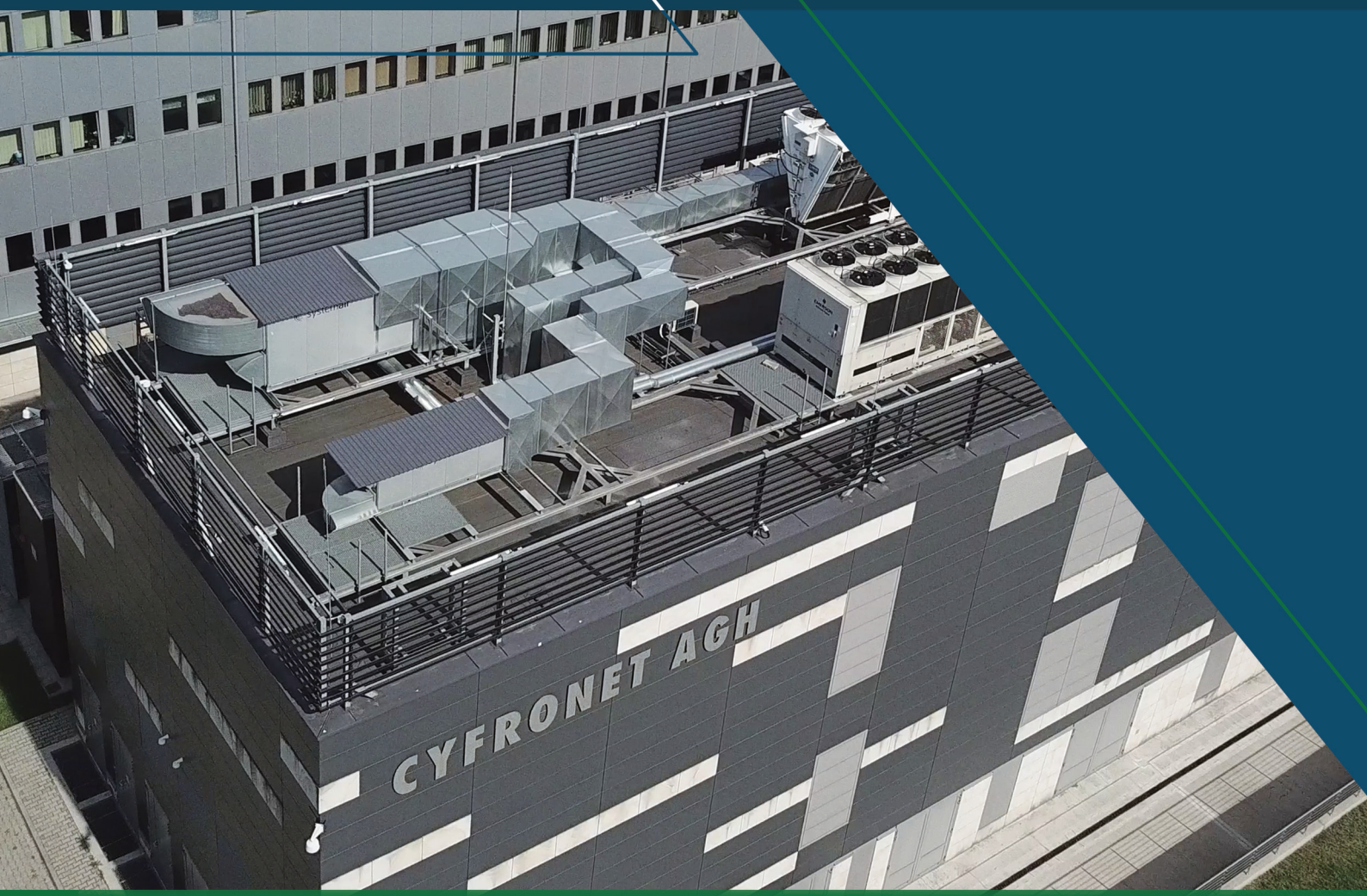
Academic Computer Centre

CYFRONET AGH



2019

ACC CYFRONET AGH is a leading unit empowered by the Committee for Scientific Research to develop and manage the High-Performance Computers (HPCs) and Cracow Metropolitan Area Network (MAN). CYFRONET is the coordinator of the PLGrid Program and is recognized by the National Centre for Research and Development as a Centre of Excellence in the area of grid and cloud services.

Dear Readers!

We are entering 2019, facing the new challenges of the 21st century. The contemporary world is developing with the unusual dynamics and this especially concerns information technology. Cyfronet, established over 45 years ago, is dynamically acting to follow these changes in ICT and actively participate in creating them. These has become possible thanks to the engagement in multiple research grants, particularly those realised by international teams and consortia. The Cyfronet's Laboratories are an important part of these activities.

When Cyfronet was founded in 1973, being the first supercomputer centre in Poland, its aims, registered in the Regulation of the Minister, was to perform numeric calculations for the scientific community, cooperate in scientific research and research conduction in terms of the use of the electronic computational techniques. Since then the Centre has extremely developed and changed, but our main mission remained the same: to support the scientific community by providing the computing power of supercomputers, large storage resources and advanced software, as well as providing access to a super-fast computer network and numerous IT tools.

The important aspect of our activity is the construction and development of a nationwide, distributed, grid and cloud computing infrastructure under the PLGrid Consortium led by Cyfronet. Thanks to the resources available through the PLGrid platform, many Polish scientists conduct their research and perform analyses at a world-class level and can compete with specialists from foreign research centres. Convenient access to computing capabilities as well as several hundred programmes, applications, platforms and services, gives the possibility to quickly obtain precise results. As the effect, original scientific works are delivered, discoveries are made, new technologies and inventions are presented, and patents are granted.

Prometheus supercomputer is the significant contributor to these successes, having over 53,000 of cores, with the performance of 2,4 Pflops, which was once again classified on the TOP500 list of the world's fastest supercomputers, taking high, 103rd position in June 2018. Prometheus is still being supported by Zeus with computing power reaching a peak performance of 374 Tflops, known particularly for its heterogenic architecture organised in four partitions. The value of these hard-working "robots" is substantially increased by the highly qualified staff who also cooperates with the researchers on provisioning the best IT support tools. We try to meet the growing needs of the scientific community, so we constantly develop our competences and expand our computing infrastructure.

Our mission is to serve the scientific community. Therefore, we actively support large, international projects as well as small research groups and individual researchers. We are proud of every Users' success coming from Cyfronet's infrastructure utilisation. For the next year I wish You even more successes! In parallel, I would like to thank all our Friends and Users for cooperation and valuable advices regarding the further development of the Centre. I would also like to invite you to personal contacts, also during our annual meetings: HPC Users' Conference, Cyfronet Open Day and international cloud and grid CGW Workshop.

Yours sincerely,
Prof. Kazimierz Wiatr
Director of ACC Cyfronet AGH

DIRECTOR
Kazimierz Wiatr, Prof.
phone: +48 12 633 34 26
e-mail: Kazimierz.Wiatr@cyfronet.pl

Deputy Director for IT Infrastructure
Karol Krawentek, MSc Eng.
phone: +48 12 633 34 26
e-mail: Karol.Krawentek@cyfronet.pl

Deputy Director for Administrative Affairs
Agnieszka Szymańska, MSc
phone: +48 12 633 34 26
e-mail: Agnieszka.Szymanska@cyfronet.pl

Deputy Director for Financial Affairs
Angelika Zaleska-Walterbach, MSc
phone: +48 12 633 80 53
e-mail: Angelika.Zaleska@cyfronet.pl

Computer Networks Department – Krzysztof Gaweł, MSc (manager)
Phone: +48 12 634 10 25, e-mail: Krzysztof.Gawel@cyfronet.pl

High Performance Computers Department – Patryk Lasoń, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Patryk.Lason@cyfronet.pl

Data Storage Department – Marek Magryś, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Marek.Magrys@cyfronet.pl

HPC Software Department – Łukasz Flis, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Lukasz.Flis@cyfronet.pl

High Performance Computing Users Department – Mariusz Sterzel, PhD (manager)
Phone: +48 12 632 33 55, e-mail: Mariusz.Sterzel@cyfronet.pl

Operational Center Team – Andrzej Zemła, PhD (manager)
Phone +48 12 632 33 55, e-mail: Andrzej.Zemla@cyfronet.pl

Technical Department Nawojki – Damian Trela, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Damian.Trela@cyfronet.pl

Technical Department Pychowice – Mariusz Kula, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Mariusz.Kula@cyfronet.pl

Administrative Department – Agnieszka Szymańska, MSc (manager)
Phone: +48 12 633 34 26, e-mail: Agnieszka.Szymanska@cyfronet.pl

Financial Department – Angelika Zaleska-Walterbach, MSc (manager)
Phone: +48 12 633 80 53, e-mail: Angelika.Zaleska@cyfronet.pl

Projects Department – Aleksandra Pałuk, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Aleksandra.Paluk@cyfronet.pl

PLGrid Program and European Collaboration Team – Jacek Kitowski, Prof. (manager)
Phone: +48 12 633 34 26, e-mail: kito@agh.edu.pl

Laboratory of Computing Acceleration and Artificial Intelligence
Paweł Russek, PhD DSc Eng. (manager)
Phone: +48 12 633 34 26, e-mail: Pawel.Russek@cyfronet.pl

Laboratory of Cloud Technologies – Łukasz Dutka, PhD (manager)
Phone: +48 12 633 34 26, e-mail: Lukasz.Dutka@cyfronet.pl

Laboratory of Data Processing – Tomasz Szepieniec, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Tomasz.Szepieniec@cyfronet.pl

Laboratory of Informatics Methods in Medicine – Marian Bubak, PhD Eng. (manager)
Phone: +48 12 633 34 26, e-mail: bubak@agh.edu.pl

Laboratory of Parallel Algorithms – Bogusław Cyganek, Prof. (manager)
Phone: +48 12 632 33 55, e-mail: cyganek@agh.edu.pl

Laboratory of Visual Techniques – Jacek Przybylski, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Jacek.Przybylski@cyfronet.pl

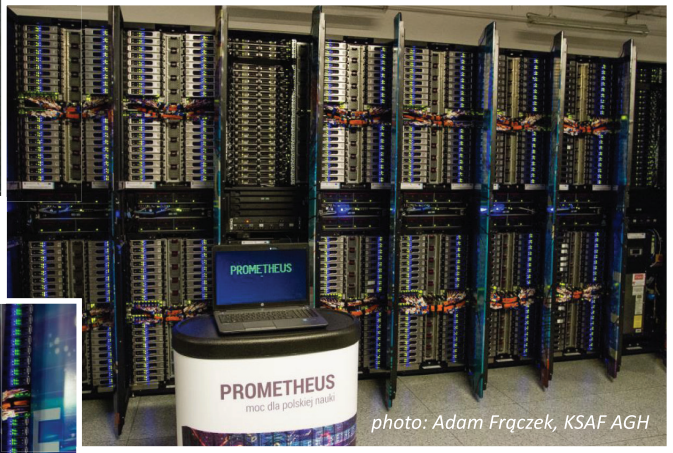


photo: Adam Frączek, KSAF AGH

photo: Krzysztof Mastalski, KSAF AGH

The Users' Council

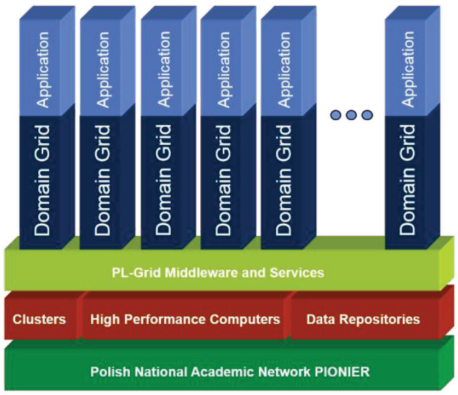
Marek Jarnicki, Prof. The Chairman	Jagiellonian University
Marek Jeżabek, Prof. Deputy Chairman	Institute of Nuclear Physics Polish Academy of Sciences
Bogdan Bochenek, MSc	Institute of Meteorology and Water Management
Andrzej Dietrich, MSc	Institute of Petroleum and Natural Gas
Jan Dudek, MSc Eng.	Cracow University of Economics
Florian Gambuś, Prof.	Agricultural University in Cracow
Jerzy R. Jaworowski, PhD Eng.	Cracow University of Technology
Paweł Kruszelnicki, MSc Eng.	Academy of Physical Education
Robert Gryboś, PhD	Jerzy Haber Institute of Catalysis and Surface Chemistry Polish Academy of Sciences
Wojciech Majka, MSc Eng.	Jagiellonian University
Paweł Mikłaszewski, MSc	Academy of Music
Roman Mazur SDB, PhD	Pontifical University of John Paul II in Cracow
Andrzej R. Pach, Prof.	AGH University of Science and Technology
Wojciech Palacz, PhD	Jagiellonian University
Bartłomiej Pokrzywka, PhD	Pedagogical University in Cracow
Marek Policht, MSc Eng.	Institute for Advanced Manufacturing Technologies
Tomasz Szmuc, Prof.	AGH University of Science and Technology
Łukasz Łakomski, Director	Marshal Office of the Małopolska District
Grzegorz Żych, Director	Cracow City Hall

PLGrid Program

The infrastructure and the projects

The idea of the PLGrid Program has been invented by Cyfronet as a result of vast knowledge and experience gained in variety of national and EU projects. In 2007, it constituted formally as the PLGrid Consortium consisting Interdisciplinary Centre for Mathematical and Computational Modelling in Warsaw, Poznan Supercomputing and Networking Center, Wroclaw Centre for Networking and Supercomputing, Academic Computer Centre in Gdansk and Academic Computer Centre Cyfronet AGH as the initiator and coordinator of the PLGrid Program and Consortium. The work carried out by Consortium partners led to the full-fledged distributed infrastructure for scientific computing. This infrastructure comprises not only high performance computing hardware, but also mass storage and dedicated tools for deployment of scientific applications on the available resources.

The design and construction of the PLGrid infrastructure started in the framework of the PL-Grid project (Polish Infrastructure for Supporting Computational Science in the European Research Space), in response to science needs, in which computers become more and more important. The main goal of the built infrastructure was to support scientists' investigations by integrating experimental data and results of advanced computer simulations carried out by geographically distributed research teams with use of supercomputers localised in High Performance Computing Centres. This aim was accomplished, among others, by extending the amount of computational resources in all PLGrid Consortium institutions. What is more, thanks to the PL-Grid project, in fall 2011 all Consortium partners have been spotted on TOP500 – the list of fastest world supercomputers. The same year Zeus supercomputer in Cyfronet has been located at 81st position – what gave it the first place among Polish supercomputers.



The next step of the PLGrid Program was to provide the researchers with necessary IT support through preparation of the specific computing environments, i.e., services and software as well as helping users in planning, running and analysing complex scientific experiments. Preparation of dedicated computing environments, so called domain grids, tailored to the needs of 13 different groups of scientists, was the most important task of PL-Grid follow-up – implemented within the PLGrid Plus project (Domain-oriented services and resources of Polish Infrastructure for Supporting Computational Science in European Research Space).

Adaptation of the infrastructure to the needs of scientists brought by domain grids was a great success of the PLGrid Plus project. Therefore, these activities have been further extended by the "New generation domain-specific services in the PL-Grid Infrastructure for Polish Science" project.

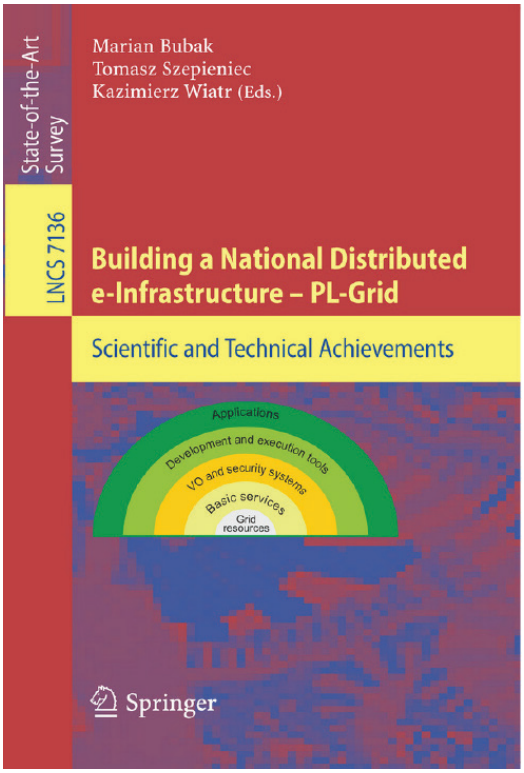
In the PLGrid NG project, the domain-specific grids were developed for several other groups of scientists, representing fourteen research fields (in total, in the two projects, IT support tools were built for 27 scientific disciplines).

However, the PLGrid Program did not stop on development of domain-oriented solutions only. Thanks to longstanding involvement in the development of grid computing infrastructures, Cyfronet is now recognized as a Centre of Excellence in the area of cloud and grid services – an achievement reflected by the new large-scale scientific grant named Distributed Computer and Data Infrastructure Centre of Excellence – PLGrid Core. This grant represented the next step in the development of the PLGrid Program and extension of the infrastructure towards Cloud Computing and handling big data calculations. It aimed not only at extension of hardware and software portfolio, but also dedicated accompanying facilities. One of them – a new backup Data Center built in separate geographical location highly increased security of scientific data sets.

It is worth noting that on the November 2015 edition of TOP500 the Prometheus supercomputer, deployed at Cyfronet in 2015 in the framework of PLGrid Core, took the 38th position, the highest so far for supercomputers deployed in Poland!

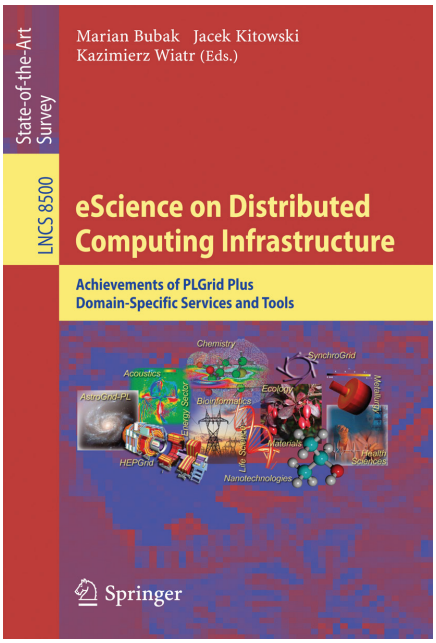
At the beginning of 2018, more than 5 PFLOPS of computing power and more than 60 PB of disk storage were available within the infrastructure. In addition, many tools supporting organization of computational experiments, designing and running applications, computationally supporting research and results' visualization were implemented in the infrastructure. Furthermore, the Consortium introduced a new service – Cloud Computing.

All the projects of the PLGrid Program have been co-funded by the European Regional Development Fund as part of the Innovative Economy program. ACC Cyfronet AGH has the honour to be their responsible coordinator. Vast range of services contributes to increase of cooperation between Polish scientists and international groups of specialists from many different scientific domains – also humanities and social sciences. The essential fact is that anyone who is performing scientific research can be the user of the infrastructure. Access to huge computational power, large storage resources and sophisticated services on a global level is free to Polish researchers and all those engaged in scientific activities associated with the university or research institute in Poland. All one has to do is to create an account via the PLGrid Portal.



Domain-specific grids in the PLGrid infrastructure

The PLGrid infrastructure, established through the Cyfronet's initiative, offers a uniform access to resources of all five Polish High-Performance Computing centres. Unification takes place at many levels, ranging from a user's single login and password across the infrastructure, to the access to scientific applications. Sometimes, however, the use of modern computing systems, services and tools of the e-infrastructure becomes relatively difficult for researchers. Basic infrastructure services are often insufficient to conduct scientific research, particularly in the context of large international consortia. In such situations, users need both assistance and close collaboration with service providers.



Therefore, within the PLGrid Plus project (2011-2015), the PLGrid infrastructure has been extended with specific environments, solutions and services, developed according to the identified needs of 13 pilot groups of scientists. The main aim of the project was to lower the barriers required for researchers to use the infrastructure, and, thus, attract new communities of users, who need the computational power and large disk space of supercomputers, but have no or little skills in using it. To enable and facilitate development of domain-specific environments, the project relied on a broad cooperation with representatives of various disciplines, often grouped in domain consortia.

The dedicated services hide the complexity of the underlying infrastructure and, at the same time, expose the actual functions that are important to researchers of the given domain. In this way, users are provided with exactly the functionality they need. What is more, it is exposed to them in their domain-specific manner to achieve maximum intuitiveness and usefulness.

Scientific and technical achievements of PLGrid Plus were presented in a book published in the Springer Publisher, in September 2014. The book is an important source of information for researchers, developers and system administrators, who use grid and cloud environments in their research. The book contains 36 chapters and is divided into three parts: the first one (chapters 1 to 8) provides a general overview of the work carried out in the project and a description of the current state of the PLGrid infrastructure, including new solutions in the field of security and middleware.

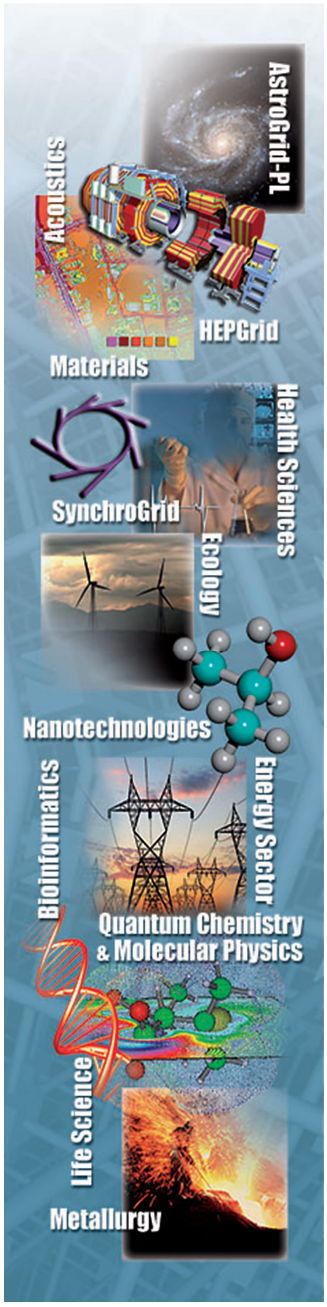
The second part (chapters 9 to 13) presents new environments and IT services that can be used by all of the previously mentioned groups of scientists. The third part (chapters 14 to 36) describes how specific environments, tools and services, prepared within the PLGrid Plus project, are used in advanced computations and computer simulations performed by different groups of researchers. These chapters present computational models, new algorithms and methods of their implementation using available tools and services.

Success of the PLGrid Plus project, in particular, the growing popularity of specialized tools and platforms prepared for the members of the first 13 strategic areas of science, led to a rapid increase in demand for related services to researchers in other fields. Therefore, the PLGrid Consortium launched the PLGrid NG project (2014-2015), whose primary objective was to implement, within the PLGrid infrastructure, several additional computing services for groups of scientists representing 14 new research fields.

New domain-specific services covered a wide range of activities: including provision of the specialized software, mechanisms of data storage and modern platforms integrated with a new type of tools and dedicated databases, which sped up research conduction as well as streamlined and automated the work of research groups.

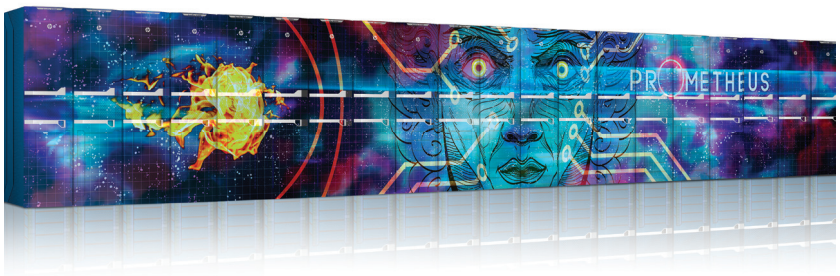
Preparation and implementation of a set of domain-specific services fit very well with the need of development of an advanced IT infrastructure designed for the implementation of modern scientific research. The well-tailored PLGrid e-infrastructure does not only fulfil researchers' needs for suitable computational resources and services, but also enables Polish scientific units collaboration with international research organizations.

Expansion of the existing computational infrastructure towards domain-specific solutions for research teams allowed more effective research conduction.



Prometheus – PetaFLOPS Computing Power

Changes in the world of science follow very quickly and affect the speed of development of IT facilities, which Cyfronet offers to scientists. Researchers' growing demands for computing power and data storage are clearly visible from the disciplines almost traditionally associated with high-performance computers: chemistry, physics, astronomy, life sciences and fields related to them. Astronomy, astrophysics and space physics are based on the one hand on data acquisition and analysis, and on the other on complex computer simulations. Biological, chemical and medical sciences as well as those mentioned above are characterized by rapid development and introduction of new, increasingly sophisticated research methods, e.g. molecular techniques based on high-performance DNA sequencing. Medicine, as a multidisciplinary field, deals with a number of time-consuming analyses, e.g., the human genome. It results in increased demand for automated collection, storage and analysis of biomedical signals and images, what in turn leads to necessity of use of the supercomputing resources in order to implement these processes. The possibility of linking together multiple unique data, i.e. the clinical, genetic as well as environmental and social data, brings many benefits, but also in this case the dedicated services are needed that can be offered only by supercomputing centers.



These are the tasks Prometheus – the most powerful Polish supercomputer – deals with. As the successor of Zeus, it has become a part of the PLGrid infrastructure and serves scientists, also within international research projects. Prometheus is used for: data results analysis, numerical simulations, (big) data processing, and advanced visualisations provision.

Prometheus consists of more than 2,200 servers based on the HP Apollo 8000 platform, combined with the super-fast InfiniBand network with 56 Gbit/s capacity. Its energy saving and high-performance Intel Haswell processors offer 53,604 cores. These are accompanied by 282 TB of DDR4 RAM and by two storage file systems of 10 PB total capacity, and 180 GB/s access speed. Prometheus has also been equipped with 144 NVIDIA Tesla GPGPUs. The theoretical performance of Prometheus is 2.4 PFLOPS (PetaFLOPS)!

Due to the innovative technology of direct liquid cooling of processors and RAM modules, Prometheus is also one of the most energy-efficient computers in its class in the world. This was achieved by using the cooling water having a temperature of 28°C. To cool down the water to such a temperature in our climate it is enough to use cheap in use dry-coolers, instead of ice water generators, consuming relatively large amounts of electricity. With use of water cooling, electronic components operate at

temperatures lower than normal, what positively affects not only the failure, but also allows to reach efficiency more than 5% higher than for a similar installation based on the classic air cooling. Furthermore, liquid cooling allowed for extremely high installation density of 144 computing servers in one rack, therefore Prometheus, weighing of more than 40 tons, covers 18 m^2 area and is placed on 20 racks only. To achieve the same computing power in case of Zeus (Prometheus' predecessor), it would have to take about 160 racks. Baribal, the predecessor of Zeus with computing power of 1.5 TFLOPS, was placed on 8 racks. To achieve the computing power of Prometheus it would take as many as 12,000 Baribal's racks.

Prometheus in numbers	
Number of computing cores	53,604
RAM	282 TB
Number of GPGPUs	144
Computing power	2.4 PFLOPS
TOP500 – the list of the world's fastest computers (June 2017 edition)	103rd position

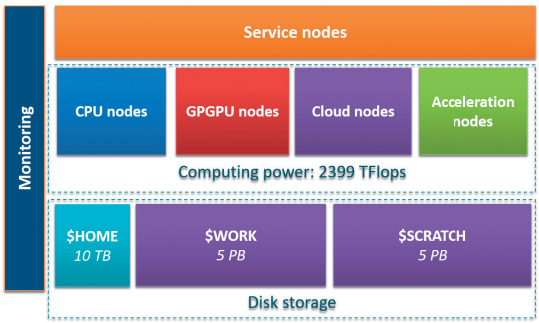
Prometheus has been installed in a high-tech computing room, exclusively adapted for its operation. The supercomputer's proper functioning is additionally supported by the accompanying infrastructure, including such systems as guaranteed power supply with an additional generator, modern air-conditioning and gas extinguishing.

Prometheus once again has been listed on the TOP500 list of the world's fastest computers (June 2018 edition) and took the 103rd position, what gave it the first place among Polish supercomputers.

Division into parts with diverse functionality, applied in the Zeus supercomputer, has been very well used by its users. Due to this fact, the Prometheus architecture is also a composite of several classes of nodes, varying in terms of architecture of computing resources and functionality:

- classical cluster of computing servers with highly efficient CPU nodes equipped with two Intel Xeon processors,
- cluster of servers equipped with graphic accelerators GPGPU NVIDIA Tesla K40 XL,
- set of servers designed for Cloud computing,
- acceleration partition with a set of devices supporting the Prometheus configuration with several types of accelerators (including GPGPU NVIDIA K80, Intel Xeon Phi 7120P, and Nallatech FPGA cards).

Prometheus architecture



Thanks to Prometheus users have received more than six times greater opportunities compared to the ones offered by Zeus. Much more efficient processors, faster network, and a greater amount of memory of Prometheus enable to perform calculations on a scale impossible to achieve using previous Cyfronet's resources.

Year	No. of Jobs	CPU time in years
2015	1 099 822	5 811
2016	3 080 543	21 239
2017	5 032 438	36 600

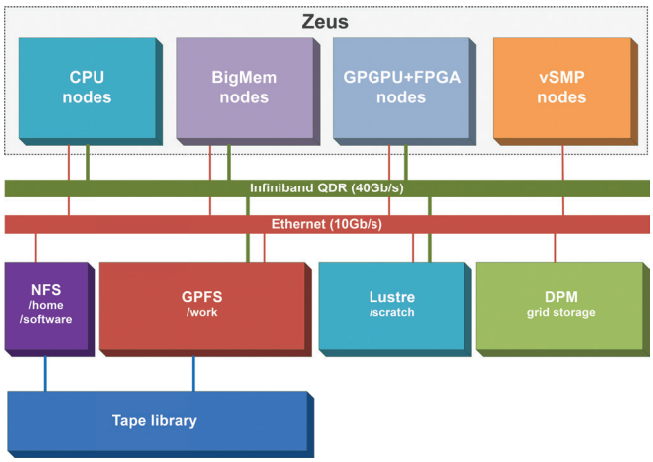
ZEUS – over 55 000 CPU-years

Cyfronet operates one of the fastest supercomputing systems in Poland, named Zeus. It currently provides 374 TFLOPS of theoretical performance, over 25,468 CPU cores and 200 GPGPUs. All this, equipped with 60 TB of RAM and 2.3 PB of disk storage supports the computations of scientific communities.

The Zeus supercomputer was launched in 2008. In 2008-2015, it was noted twice a year on the TOP500 – the list of the world's fastest computers. Four of these locations were on TOP100 subset, with 81 – the highest noted spot. Last time Zeus was noted at 387th location, in the November 2015 edition of the TOP500, announced at the Supercomputing'15 conference in Austin, USA.

The architecture

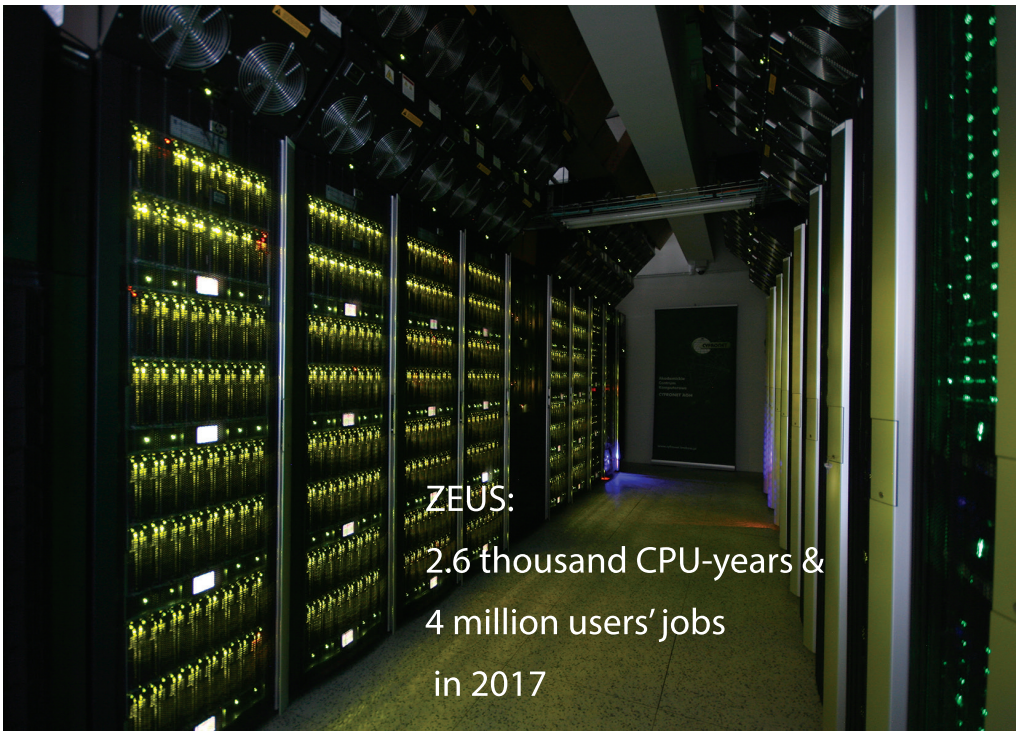
Zeus is a heterogeneous computing cluster. It constitutes of four classes of nodes, varying in terms of architecture of computing resources, specifically tailored to the requirements of the scientific communities. The Zeus architecture is a composite of four partitions:



- classical cluster of computing servers with highly efficient CPU nodes equipped with two Intel Xeon processors and 16-24 GB of memory per node,
- cluster of servers with large amount of memory – "fat nodes" with four AMD Opteron processors and 256 GB of memory per node,
- set of servers equipped with GPGPU accelerators (Intel Xeon processors as well as NVIDIA M2050 and NVIDIA M2090 cards) and FPGA accelerators (Pico Computing M-503 modules with Xilinx Virtex-6 LX240T),
- "virtual" SMP computer with a large, shared memory, using vSMP software of the ScaleMP company the nodes with Intel Xeon processors connected with a specialized virtual machine hypervisor, which allows for booting up the machines up to 768 cores and 6 TB of memory.

Diversification of the node types gives a possibility to fit users' applications to the hardware, which matches at best their characteristics and special requirements. For example, the classic CPU node group is dominated by serial and parallel (MPI) jobs, while the second one is great for large memory jobs. The GPU nodes allow some applications to benefit from GPGPU accelerators and the vSMP nodes give a possibility to run huge memory jobs or scale applications, which do not use any

inter-node communication library, like MPI, for parallelism. It is worth noting that Zeus-vSMP was the Europe's biggest installation of this type when launched!



The users

Since it was launched, the Zeus cluster has been serving the whole scientific community from Poland. In 2017 alone Zeus performed over 4 million tasks serving scientists from many universities and research institutes. Majority of tasks were executed on more than one processor and this trend is constantly increasing. Typical user computations request even several thousand of cores just for one job!

Year	No. of Jobs	CPU time in years
2008	603 525	207
2009	2 227 804	876
2010	4 009 049	990
2011	7 557 817	5 052
2012	8 126 522	7 923
2013	7 932 978	11 016
2014	7 694 224	12 980
2015	6 405 941	10 141
2016	4 668 134	3 414
2017	4 034 454	2 632

Supercomputers usage

Prometheus and Zeus are part of the European cloud and grid infrastructure under the European Grid Infrastructure (EGI). At the same time, Prometheus and Zeus are also important supercomputers in the PLGrid nationwide computing infrastructure – the platform for conducting in silico research and enabling calculations with use of high-performance computers, also within the cloud and grid architecture.

Via the PLGrid infrastructure scientists can get access to the Prometheus and Zeus resources. Dedicated computing environments, so-called domain grids, and specialised IT platforms enable conduction of increasingly complex research problems. The research portfolio carried out with the help of Zeus and, recently, Prometheus is quite reach. It includes:

- modeling of a high-frequency distributed ultrasonic wave in the skin tissue,
- study of new RNA forms in human glioblastomas,
- preparation of a LINAC device simulator for cancer therapy,
- study of wood-glass composite beams subjected to repeatedly variable loads,
- analysis of two-component dark matter in the MRSSM model,
- simulations of operation of thermoelectric modules and generators,
- measurement of the neutron electric dipole moment,
- numerical analysis of the reaction of selected engineering structures for kinematic extortion of mining origin.

A wide range of research topics is evidence of constantly increasing number of scientists, who are aware of advantages of supercomputers like Zeus or Prometheus. With the help of these powerful supercomputers one can get the final results of huge simulations many, many times faster, compared to the case of an ordinary, desktop computer. Supercomputers enable to significantly reduce time of computations that using a single computer would often take many years (in specific cases more than 150, 700 or even 1000 years). Here they may be usually performed within a few days. What is important, Cyfronet users can benefit from the professional support – starting from full documentation, through training, to individual consultations with experts.

In addition to individual scientists and small research groups, even international consortia carry out calculations from many different scientific disciplines with the help of supercomputers – of course with the participation of Polish scientists. Scientific computations do not include simulations only. Computing power is utilised by Polish researchers also within international projects, including experiments like CTA, LOFAR, EPOS, Large Hadron Collider in CERN and the recently discovered gravitational waves in LIGO and VIRGO detectors.

Obviously, even the highest positions in the TOP500 list, or the latest technologies used to build high-performance computers do not fully reflect the importance of this kind of computing resources for

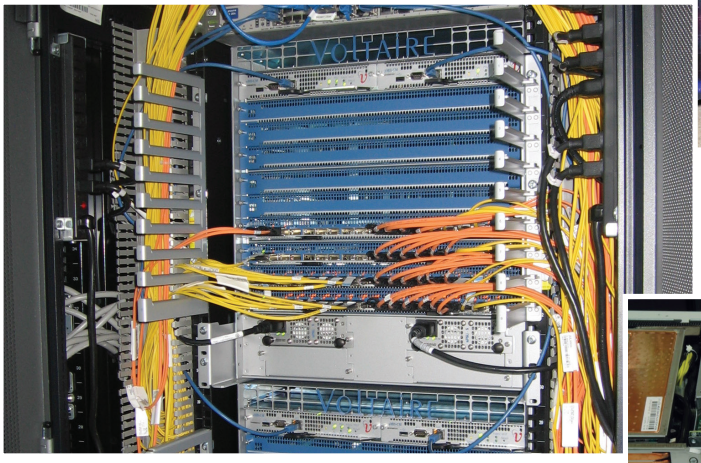
the Polish scientific community. The usefulness of supercomputers provided by ACC Cyfronet AGH as a tool for conducting research is best evidenced by statistical data on their use.

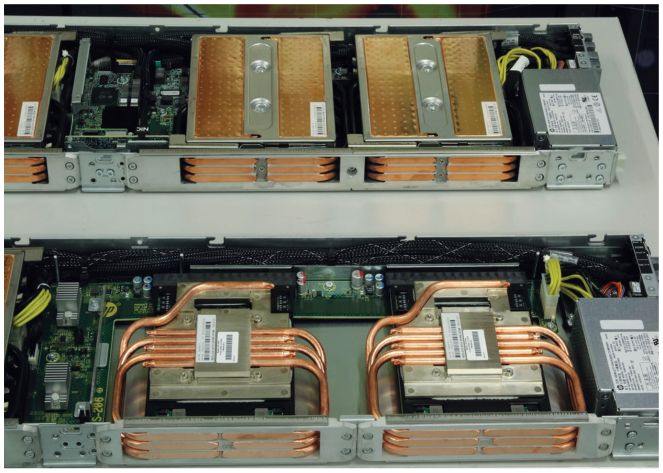
The table presents the aggregated key data on the number of computational tasks and their duration, performed by Cyfronet for other units.

It is worth mentioning that huge users' demands for computing power and space for data storage would not be fulfilled without continuous extension of computing resources and disk storage. Therefore, we carefully analyse users' suggestions and statistical data related to carried out computations together with world's trends in computing.

Year	No. of Jobs	CPU time in years
Zeus supercomputer		
2008	603 525	207
2009	2 227 804	876
2010	4 009 049	990
2011	7 557 817	5 052
2012	8 126 522	7 923
2013	7 932 978	11 016
2014	7 694 224	12 980
2015	6 405 941	10 141
2016	4 668 134	3 414
2017	4 034 454	2 632
Prometheus supercomputer		
2015	1 099 822	5 811
2016	3 080 543	21 239
2017	5 032 438	36 600
Zeus and Prometheus altogether		
2015	7 505 763	15 952
2016	7 748 677	24 653
2017	9 066 892	39 232


PROMETHEUS


VOLTAIRE
VOLTAIRE



Data storage

Even the most advanced, highly efficient HPC systems become ineffective without an appropriate data storage system. Only the proper teaming of computing infrastructure with the right selection of storage solutions can assure the best quality of services provided to users. The scale of problems in this area increases with the complexity and the efficiency of high performance computers. At present, disk storage systems attached to Cyfronet's supercomputers store over 500 000 000 data files (with the file sizes up to several terabytes). A wide variety of research conducted on the Centre's resources requires not only diverse configuration of key Cyfronet's computers, but also an efficient, dedicated storage system.

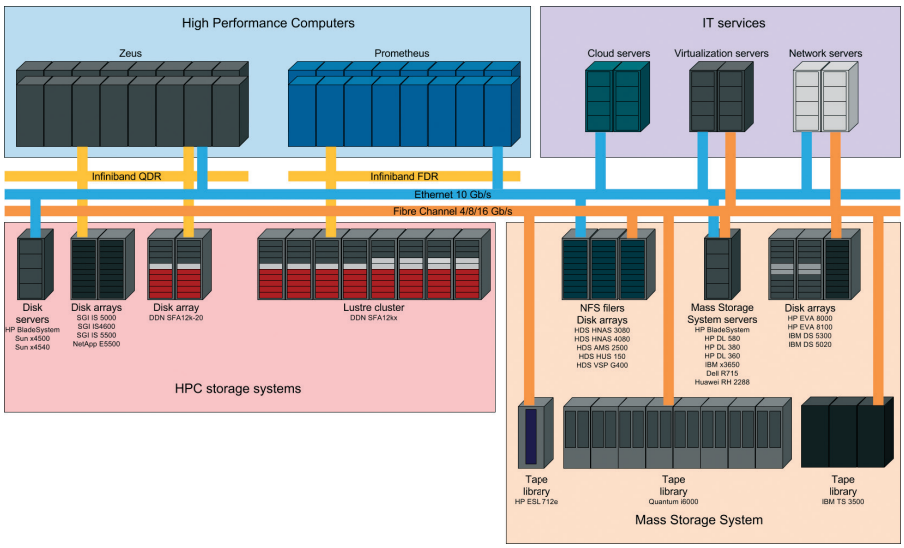
The most fundamental is the one used for keeping users' home directories. In this case all the crucial elements provide a very high level of availability and data security, which are supported by mechanisms such as snapshots and backups to external tape libraries. Zeus and Prometheus (the two main supercomputers of the Centre) offer such functionality through using specialized HNAS file servers (so called filers), produced by Hitachi Data Systems. These servers base on a hardware implementation of the NFS protocol and provide very high performance and high availability of the file systems. HNAS filers are coupled with Hitachi Data Systems AMS 2500 and HUS 150 disk arrays, used as repositories of physical disk space. These devices also provide extremely high levels of security and performance, fitted to the specific characteristics of the data stored in home directories.

Another type of storage space used in supercomputers is the scratch space, in which the crucial factor is speed. To address this requirement, Cyfronet uses Lustre distributed file system, which is capable to scale both space and performance by aggregating storage capacity of many servers. Moreover, throughput and/or capacity can be easily increased by adding more servers dynamically, without interrupting user computations. Nowadays, all Cyfronet's supercomputers can use scratch spaces based on Lustre. In Zeus case, it is the file system with almost 600 TB capacity and 12 GB/s read/write bandwidth. Prometheus' scratch has enormous capacity of 5 PB and 120 GB/s read/write bandwidth. For even more demanding disk access requirements it is possible to use a super-fast RAM-disk provided by the vSMP partition of the Zeus supercomputer.

However, the major part of Cyfronet's storage resources is dedicated to the needs of users of domain-specific services developed in the PLGrid program. The PLGrid infrastructure provides a dedicated workspace for groups in domain grid environments – the functionality essential for enabling cooperation of scientists from geographically distributed locations. Zeus provides almost 200 TB of such disk space with the use of HNAS filers and the NFS protocol. Prometheus offers similar functionality with higher performance, using the Lustre file system. The maximum capacity of the /archive resource in this supercomputer reaches 5 PB and the total rate of read/write operations attains 60 GB/s.

A special case of mass storage are the resources for large projects and international collaborations, in which Cyfronet takes part, such as WLCG (Worldwide LHC Computing Grid), which stores and analyzes the data coming out of the LHC detector in CERN, or CTA (Cherenkov Telescope Array). Such projects demand high volumes of disk space available by a set of specialized protocols, such as SRM, xroot or GridFTP. Cyfronet provides such space with the use of the DPM (Disk Pool Manager) instances and dedicated networks, such as LHCone. Total amount of disk space provided by these services exceeds 1 PB. The overall Data storage space exceeds 45 PB.

Comprehensive infrastructure of efficient and safe storage of digital data



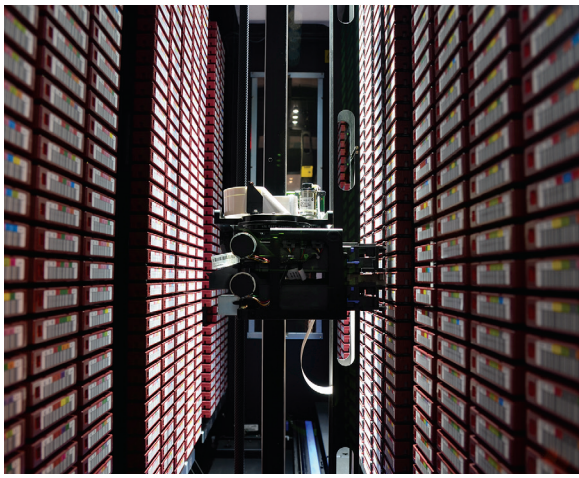
The currently observed phenomenon of the rapidly growing amount of digital information also applies to the scientific community. Access to very efficient supercomputers enables performing analyses of large-scale research problems, what results in generation of huge data sets. They require a completely new approach to information processing and storing. This problem, being currently one of the most important challenges of the modern digital world, is described by the concept of BigData. Also in ACC Cyfronet AGH there is clearly visible correlation between the growing expectations concerning available capacity, speed and additional functionalities of storage resources, caused by offering more efficient computing systems.

Extremely diverse characteristics of research topics, achieved through computational tasks performed in the infrastructure provided by the Centre, is reflected in the complexity of the architecture of the Data Storage System – the main mass storage platform for High-Performance Computers. This system is composed of following elements:

- the SAN network – the efficient and highly available network dedicated to communication among devices within the Data Storage System, and clients using shared resources or services,
- disk arrays and servers of various types, offering the storage space for the users' data – starting from fast, but expensive and less capacious solutions, and ending with the devices with large storage capacity and relatively cheap, but with limited efficiency,
- service servers, with specialised tools and virtualisation software, providing users with functionalities such as automatic backup and archival, hierarchical data storage systems, high-performance hardware file platforms or distributed network file systems,
- tape libraries and specialised software for storing users' critical data on magnetic media,
- additional infrastructure, e.g. Ethernet, Infiniband as well as solutions supporting IT infrastructure management and enabling secure storage of magnetic media.

At present, the total storage capacity of Cyfronet disk resources is 21 PB.

Backup-archiving services in detail



ACC Cyfronet AGH provides its users with a wide portfolio of services related to securing information stored in a digital form. In addition to advanced technological solutions such as communication networks dedicated to storage systems, modern disk arrays or hardware file servers, the Centre also performs conventional backup-archiving services, based on magnetic media. Contrary to the expectations of the inevitable end of solutions using data storage on magnetic tapes, this technology is constantly evolving, and offers in successive generations not only the increasing capacity of the media, but also significantly better capacities and mechanisms supporting the safety and effectiveness of the information storage (e.g. data encrypting and compressing algorithms, which are embedded in the tape drives).

Cyfronet has currently three tape libraries having in total 6 thousand slots for LTO magnetic tape drives and 36 drives of the III, IV, V and VI generation. A single LTO-6 magnetic medium has a physical capacity of 2.5 TB and allows recording at the speed up to 160 MB/s, which theoretically allows the storage of almost 15 PB of uncompressed data in tape libraries. Described resources are used for performing current backup and archive of important information resources of the Centre's users.

Backup is performed on the active data – that might be currently in use – through a replication process from the source location to a separate, isolated destination. The ideal backup procedure ensures consistency of the source and backup data, both at the level of a single object (a file located on a hard drive), and in the case of complex IT systems, such as database or mail servers as well as virtual environments. Physically, the cloning process is usually done by copying the source data from the backup client disk to disk/tape resources of the target backup server, using dedicated or shared access medium, such as Ethernet or SAN.

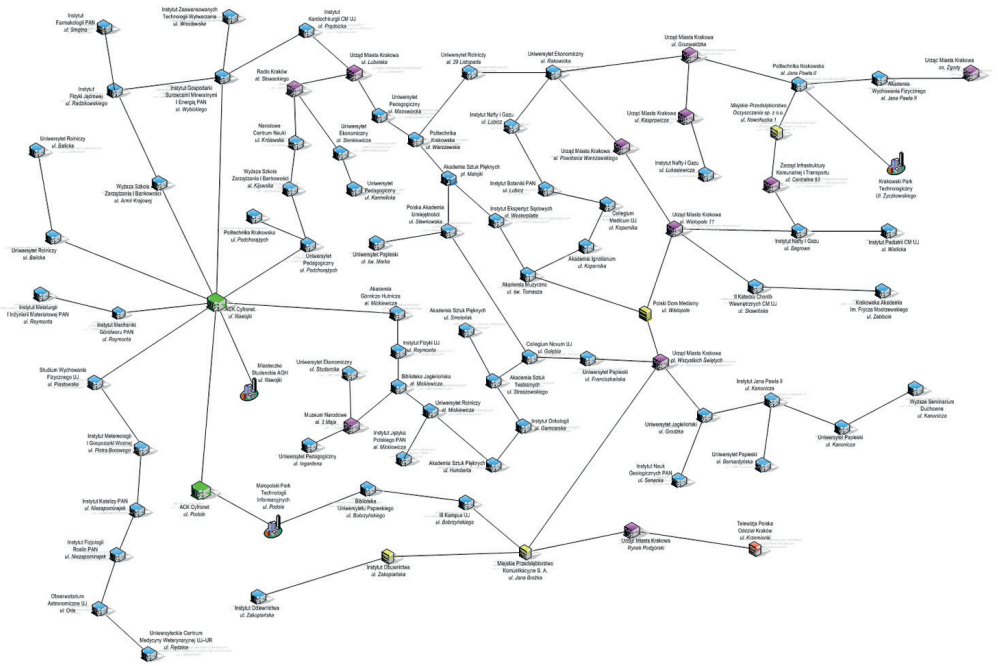
The purpose of an archive is to ensure security of unused data and to release occupied storage resources. In contrast to the backup, the archive is performed once, by the migration of the data from the source location to the destination. Since this is a migration process, the data from the source is removed, allowing the release of the disk resource.

ACC Cyfronet AGH provides a wide range of backup services, addressed directly to users, and operating without their interaction. Among those at the disposal of users, there are ones based on FTP, NFS and SCP network protocols, acting within the dedicated backup servers. These machines provide backup solutions for users, allowing them to direct access to the backup data. It is up to users to decide which data they treat as a backup and which as archives.

For the special cases Cyfronet offers users a dedicated backup-archive service called the Universal Archiving. Within this service, the user is given a dedicated disk space, protected at many levels. User's data in this case is protected by a distributed disk array equipped with disk resources protected by RAID-6 level functionality and additionally secured by a system of hierarchical data storage HSM. To advance the data safety even more, users' backup data are additionally protected by geographical data replication to the associated units. Last but not least, users of Universal Archiving system can further increase their data safety by encrypting their data with use of certificates. At present, the total storage capacity of Cyfronet tape resources exceeds 25 PB.

Metropolitan Area Network

One of the major characteristics of the present science is complexity of research challenges, including their multidisciplinary character, use of heterogeneous models, resources and massive amount of data produced by a variety of sources. Research is not performed by a small group of scientists anymore, but by international consortia. In order to bind those usually geographically distributed resources together, fast and reliable network connectivity is essential. Therefore, one of the principal tasks of the ACC Cyfronet AGH is development and maintenance of the Metropolitan Area Network (MAN) to achieve its availability 24/7.



Metropolitan Area Network of Krakow

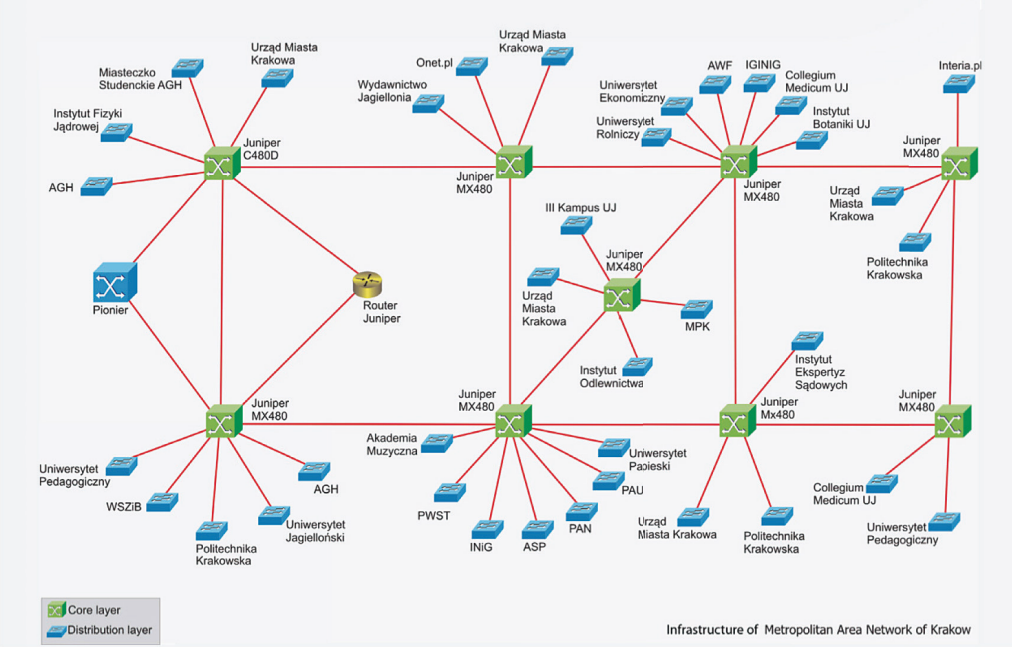
Main characteristics of MAN

It is not possible to attain high network availability without its constant development and adjustment to the needs of users. The length of dedicated fiber-optic links reached this year almost 200 km. The core links of the network are located in the Old Town area and reach the academic campus of AGH University of Science and Technology. Furthermore, the network covers also Bronowice, Krowodrza, Czyżyny and Nowa Huta zones. Recent expansion of the network included such distant research

centres like Prokocim, Borek Fałęcki and the 3rd campus of the Jagiellonian University in Pychowice. Development of the core backbone includes also other directions, up to the borders of Kraków.

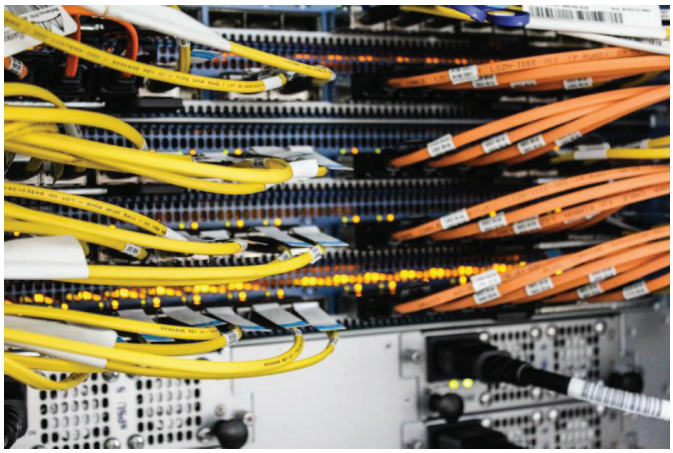
The core data link layers are implemented using top quality equipment with 1 and 10 Gb Ethernet technologies. Each of the backbone networks switches is connected with at least two and sometimes even three neighbours for automatic and transparent recovery in case of a failure of any network device or link. Our users can obtain fiber-optic connectivity to the network via 10/100/1000 Mbps, 1 Gbps or even 10 Gbps Ethernet cables as well as through traditional modem uplinks.

The Metropolitan Area Network is directly connected to Warsaw, Katowice, Bielsko-Biała and Rzeszów through the PIONIER network. Currently the links can serve up to 2x10 Gbps capacity. High Performance Computing centres in Poland (Gdańsk, Kraków, Poznań, Warsaw and Wrocław) are integrated with links of 2x100 Gbps capacity. The PIONIER network enables also communication with major national and foreign computing centres. International connectivity is achieved through the GEANT scientific network with 100 Mbps capacity. In addition, the reserve connection with 2500 Mbps capacity is established to the International Carrier TeliaSonera Poland network.



Network services provided to the users

From the beginning of the Polish Internet (mid 1991 ACC Cyfronet AGH has been actively participating in the development of the telecommunications infrastructure and, what is very important, the wide range of Web services. Those include:



- e-mail accessed via SMTP protocol or web interface http://poczta.cyfronet.pl,
- news: discussion groups covering all areas of interest – from highly specialized scientific to general-purpose boards (for sports, pastimes, art, etc.) The news server operated by ACC Cyfronet AGH registers over 100 000 new messages each day,
- www: CYFRONET operates a set of web sites, which in addition to news from the world of science, present information on the culture, sights, tourism, communications and many other fields,
- ftp: CYFRONET mirrors major international software archives, providing shareware and freeware applications for MS Windows and UNIX and LINUX systems. The establishment of this service has significantly reduced the traffic on CYFRONET's international links while at the same time enabling faster downloads of software for users of the Krakow MAN,
- eduroam: provides the academic network access at all locations on eduroam on the world with a single authorized account, providing at all locations the same way as access to the network at the parent unit,
- box: a network drive (http://box.cyfronet.pl) allowing file exchange and synchronisation. The drive can be also accessed from mobile devices via dedicated application.

Network services in numbers	
Number of e-mails	> 17 000 000
Number of e-mail server sesions	> 40 000 000
Number of news groups	> 5 800
Daily number of news messages	> 100 000
Annual number of news messages	> 36 000 000

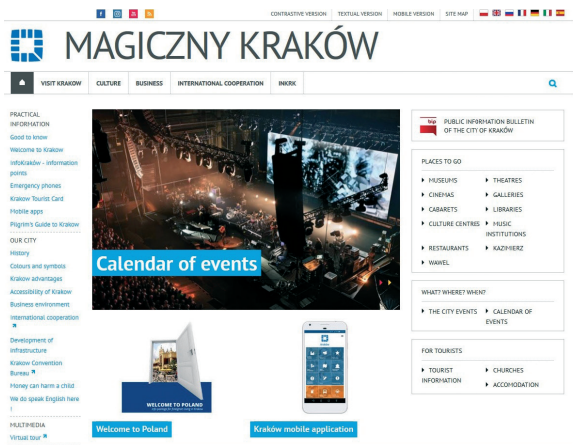
Portals and mobile applications

The Centre does not limit its activities to the scientific areas only – it also contributes to the development of the information society. The Web server at ACC Cyfronet AGH serves as an Internet hub for the entire Kraków scientific community. The Centre continues to develop and extend its Web portal, which has gained substantial popularity over the years.

Cooperation with Kraków authorities is of particular importance for the Centre. The agreement between the Municipality of Kraków and CYFRONET, regarding the promotion of the City has resulted in the creation of an up-to-date portal. Aside scientific information the portal introduces its readers to the culture, historic sites, tourism, local transit and many other aspects of life in Kraków.

In collaboration with the City Hall, the Centre has been developing and running the Internet Bulletin for Public Information in the Kraków Region. In 2005 this collaboration was extended in order to provide content services for municipal units, libraries, schools, etc.

In 2007, the "Magical Kraków" web portal – www.krakow.pl has been nominated for the World Summit Award as the best e-Government service in Poland. The mobile version of the portal was awarded at the conference Mobile Trends, Mobile in 2012 as the best city mobile web site in Poland.



Cooperation with the City Hall explores also the area of mobile devices. CYFRONET has developed – among others – a mobile application "Kraków.pl". The app can be used as a Kraków city guide, a source of important information like phone numbers, info points, consulates or pharmacies. The most important part of this app is the ability to check all those places on an offline map. Our app is available in Polish, English and Spanish.

The Cloud Computing in PLGrid

The PLGrid infrastructure has been designed with particular focus on scientists and their needs. Its character allows for easy adaptation to, even sophisticated, research challenges performed by different groups of scientists – from small research teams up to international consortia of researchers. To fulfil their requirements, in addition to typical computing and storage platforms, we offer the PLGrid Cloud Computing Platform.

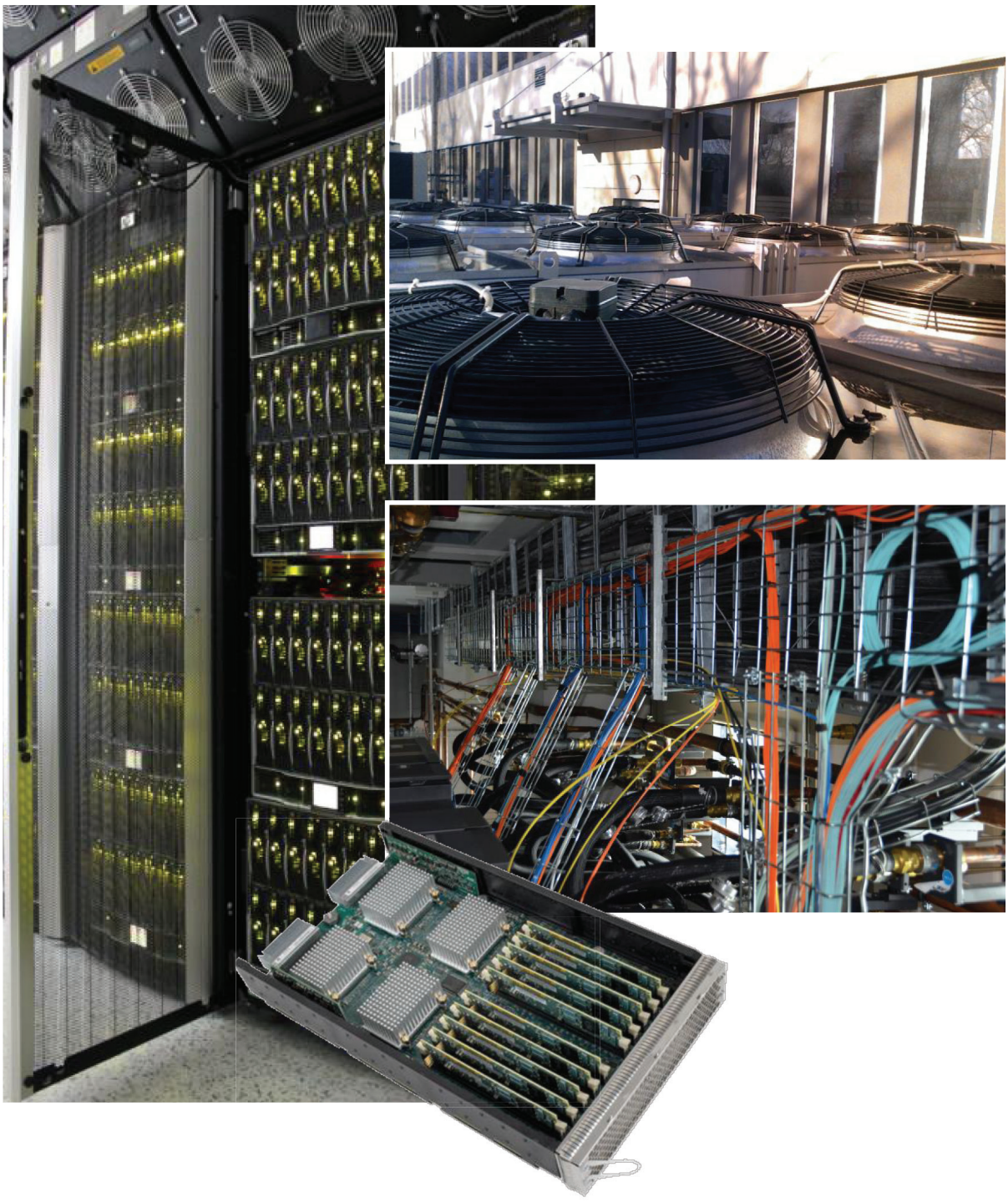
– Up to now, the PLGrid infrastructure has been providing a set of well-defined environments with computing and storage resources. The cloud platform is not just an extension of them. We foresee it as a new quality level of conducting research – says Kazimierz Wiatr, the Centre's Director. – A user can easily connect to a requested set of virtual machines (VM), with full access rights to the operating system. To achieve high security, all the VMs operate in a dedicated, local area network. Particular services can be accessed from all over the world, easing cooperation between scientists – adds Director.

There are several advantages of the cloud computing we would focus in particular:

- The Cloud increases elasticity of research, as scientists can tune the virtual machines to their specific needs. Up to now, to set-up a "virtual laboratory" solving some specific scientific problem, some help of PLGrid experts was needed. Now, each scientist can create and easily extend such virtual laboratory alone.
- The catalogue of VMs offered by PLGrid contains many OSes. Thanks to this, users can run their software applications with Operating Systems other than Scientific Linux, including Windows or other Linux OSes.
- With Cloud, it is easy to build and put in operation a test environment. This feature is very convenient for scientists developing their own software. Any test task can be then easily performed and its results analysed.
- It is possible to maintain a communication with already executed computing job. In addition, every virtual machine can be easily duplicated, even in thousands of copies or more. A start of a new VM takes just around 30 seconds.
- The Cloud platform is also the best and in many cases the only solution for running jobs with legacy software packages. In a secure LAN environment even old, deprecated operating systems can be used. This feature is also a solution for dispersed international groups using variety of different packages for their research. Every group can run their own computations and easily share their results with others.

– The Cloud Computing in PLGrid and Cyfronet is an innovative solution on a European scale. We have a strong belief it will bring a new quality level for research conducted by our users – concludes Prof. Wiatr.

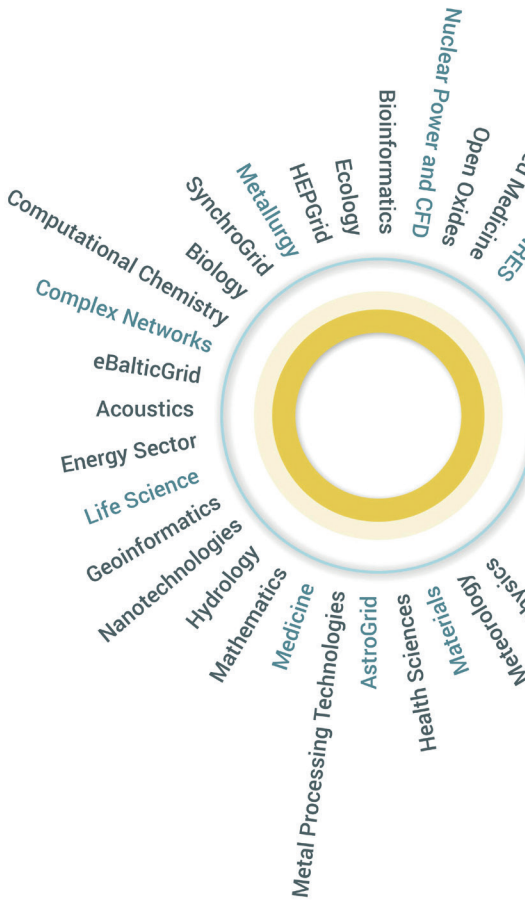
Currently 200+ various types of VMs are utilised on Cyfronet resources.



Advanced computing platforms and domain-specific services

Among the scientists conducting research with use of high-performance computers and large storage resources there is a need for different types of interaction with a computer or with the infrastructure. To address these needs Cyfronet provides a number of advanced IT platforms and dedicated services that hide the complexity of the underlying IT infrastructure and, at the same time, provide the functionalities important from the point of view of scientists from the particular field, precisely tailored to their needs.

Together with computing infrastructure we provide a selection of tools, which enable researchers to perform complex, large-scale experiments and manage their results in an easy way. The efficiency of the performed analyses and the safety of their associated data are guaranteed by appropriate IT solutions, benefitting from the extensive experience of Cyfronet's developers. The platforms have been successfully applied in the PLGrid Program for domain specific grids. As we mentioned before we have prepared more than 70 tools, platforms and services gathered into 27 scientific domains dedicated for important scientific topics and strategic fields of Polish science. All those services are provisioned in the framework of the PLGrid infrastructure, allowing Polish scientists and their foreign collaborators to access it in a convenient manner.



Among others, at the Centre we offer advanced tools and graphical interfaces that enable construction of dedicated environments for scientific research, building application portals, conducting virtual experiments, visualization of calculations' results, executing complex scenarios with parallel tasks, as well as supporting uniform and efficient access to data. All of these services are important support for researchers, as they have an impact on improving and, where possible, automating the work of research groups, what greatly accelerates obtaining research results. On subsequent pages we will learn about capabilities of selected services.

Invitation to cooperate

We are looking for people interested in development of domain-specific services. We also offer support in scientific research.

We encourage scientists to send us their program codes for the compilation by the experts at the Centre. After installation, we provide assistance in their effective use. We also enable the use of scientific software licenses held by research groups.

Computational resources

ACC Cyfronet AGH provides mature computing infrastructure for Polish science based on five main pillars. Furthermore, complex support and training are available for the users.

Computing resources

Prometheus and Zeus supercomputers provide:
2,75+ PetaFLOPS
78 000+ cores
300+ GPGPUs
340+ TB RAM

Storage

21 PB of disk and 25 PB of tape storage space and fast scratch Lustre filesystems enable big data processing and analyses.

Scientific software

Vast portfolio of tools, libraries and scientific applications for research in various fields of science.

Tools for scientific collaboration

Tools and services such as Stash Git repositories server and JIRA issue & project tracking solution ease scientific projects coordination and communication between researchers.

Computational cloud

Cyfronet's PaaS based on OpenStack provides elastic solution for computational environment which can be easily adapted to researchers' needs.

ONEDATA

Onedata is a global data management system, which provides transparent access to data stored on distributed storage resource managed by multiple providers. Onedata can scale to meet the needs of small user communities or large federations of users and storage providers, making it a perfect solution for large research initiatives, long-tail of science as well as for commercial purposes. Onedata allows users to rely on a single solution for managing their personal as well as research data sets and access them efficiently on any machine, from personal laptop as well from a Cloud virtual machine.

Onedata provides a unique federation system based on zones, which enables storage providers to organize into trusted federations and allows users to easily request storage resources from providers within a zone.

Features for users

- Unified access to data stored on heterogeneous storage systems distributed across the infrastructure. With Onedata, users can access their data from anywhere, as the system automatically replicates and transfers necessary blocks on demand.
- All data is organized into space, which can be regarded as virtual folders or volumes, accessible from any client machine via POSIX protocol.
- Easy to use web based Graphical User Interface for data access, discovery and management.
- Support for easy data sharing and collaboration with other users, while ensuring security through custom Access Control Lists and creation and management of user groups.

- Open data publishing functionality integrated into the user interface, enabling publication of prepared datasets, registration of DOI identifiers and indexing in open access portals.

Features for administrators

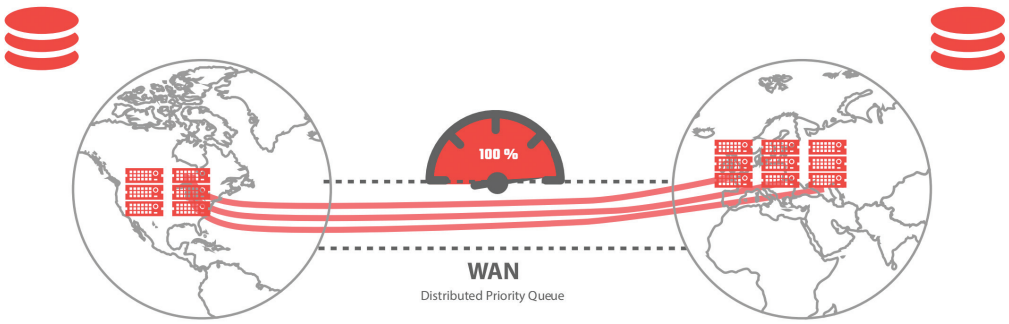
- Simple deployment based on Docker containers using a friendly command line client.
- Easy storage support for user requests based on secure tokens.
- Complex monitoring information available on all aspects of the system, accessible through REST API or directly visualized in the administration panel of the Graphical User Interface.
- Support for multiple storage backends including POSIX based storage (e.g. Lustre), S3, Ceph, OpenStack SWIFT, and GlusterFS.

Features for developers

- Easy integration with Onedata services using REST API and CDMI protocols.
- Flexible authentication and authorization of requests based on Macaroon tokens.
- Complete reference documentation of the REST API including sample clients for several programming environments.

Onedata users

Onedata is currently deployed and evaluated in several initiatives in Europe including Polish National Grid infrastructure PLGrid, INDIGO-DataCloud, EGI DataHub, Human Brain Project and Helix Nebula Science Cloud. In HBP it has proven to meet the users' hard requirements of real-time brain visualization use case.



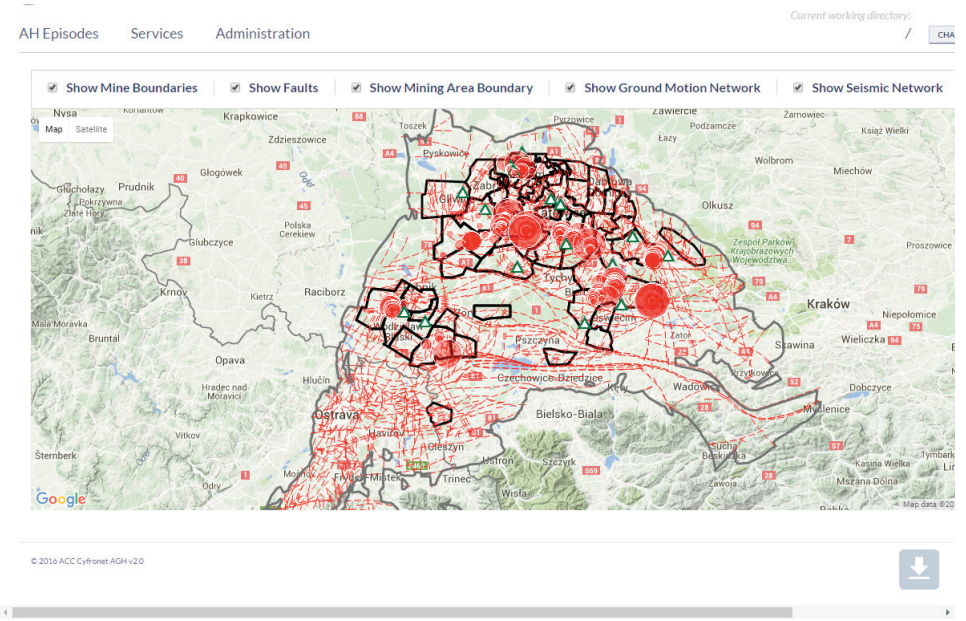
More information: https://onedata.org

InSilicoLab is a framework for building application portals, also called Science Gateways. The goal of the framework development is to create gateways that, on the one hand, expose the power of large distributed computing infrastructures to scientists, and, on the other, allow the users to conduct in silico experiments in a way that resembles their usual work.

The scientists using such an application portal can treat it as a workspace that organizes their data and allows for complex computations in a manner specific to their domain of science.

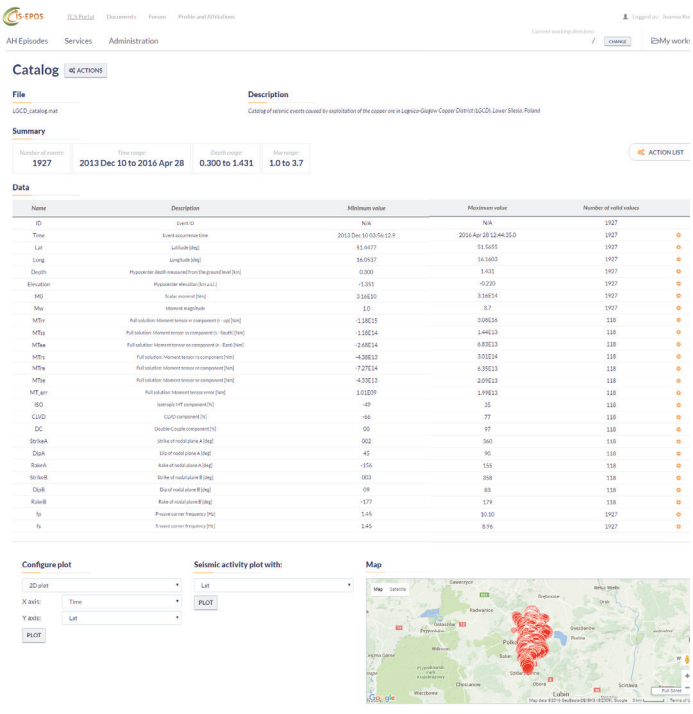
An InSilicoLab-based portal is designed as a workspace that gathers all that a researcher needs for his/her in silico experiments. This means:

- capability of organizing data that is a subject or a product of an experiment – this should include:
  - facilitating the process of preparation of input data for computations,
  - possibility of describing and categorizing the input and output data with meaningful metadata,
  - searching and browsing through all the data based on the metadata,
- seamless execution of large-scale, long-lasting data- and computation-intensive experiments.



The approach

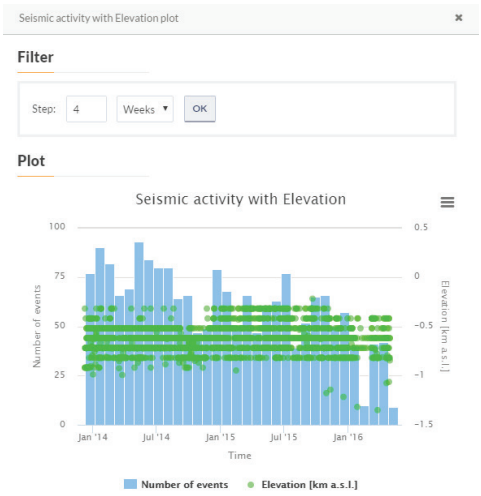
InSilicoLab is not meant to be a "Yet another engine for...", therefore, its developers has put maximum stress on the utility of the tool. This means user-friendliness, but, even more importantly, serving real scientific problems. This requires focusing on solving specific problems, rather than building a platform to solve any scientific problem, as the latter cannot be done in a universal and comprehensive way. Therefore, building a framework, which obviously is a generic solution, has to be performed in a bottom-up approach – starting from the particular problems, and building the generic tool from the common parts of the specific solutions.



Domain applications

Every gateway based on the InSilicoLab framework is tailored to a specific domain of science, or even to a class of problems in that domain. The core of the framework provides mechanisms for managing the users' data – categorizing it, describing with metadata and tracking its origin – as well as for running computations on distributed computing infrastructures. Every InSilicoLab gateway instance is built based on the core components, but is provided with data models, analysis scenarios and an interface specific to the actual domain it is created for.

http://insilicolab.cyfronet.pl

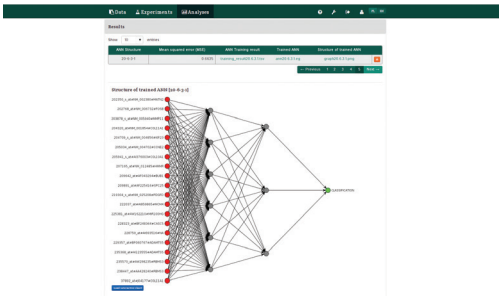


Life science and personalized medicine services

Out of disk space at your institute? Stuck waiting for alignment results? Can't easily share your results with collaborators? Or, perchance, are you looking for bioinformatics experts to assist you with managing large datasets? We can help!

We have developed a suite of tools which enables researchers representing various domains of science to perform complex large-scale experiments and manage their results. The efficiency of analysis tasks and security of scientific data are looked after by experts from the Academic Computing Centre Cyfronet AGH. Our tools reside in the PLGrid infrastructure which allows Polish scientists and their foreign collaborators to make use of vast computational services – free of charge and in a convenient manner.

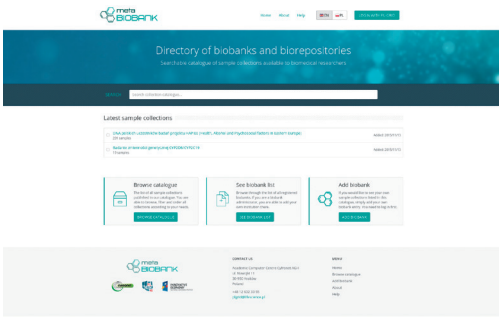
The life science toolkit includes:



DNA microarray integromics analysis platform

Designed for researchers who perform biological assays with the use of DNA microarrays. This service helps analyse gene expression data and correlate it with other clinical data sources characterising the target organisms.

https://integromics.plgrid.pl



MetaBiobank

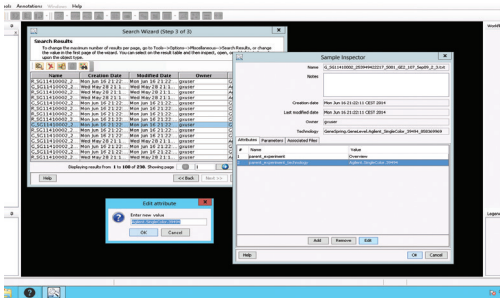
A convenient directory of biobanks and biorepositories, with searchable catalogue of sample collections available to biomedical researchers. It makes it possible to publish your own biological samples. Also supports searching for interesting samples in available biobanks and helps to establish new collaborations with other researchers. Supports ICD10 dictionary.

https://biobank.plgrid.pl

GeneSpring GX

An Agilent software package facilitating execution of statistical analyses and visualization of microarray data.

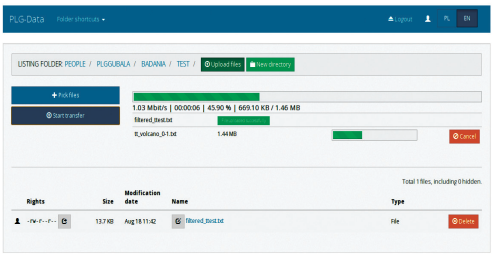
Accessible through a remote desktop interface.



PLGData

A tool for management of data stored in the PLGrid infrastructure. Comes with a user-friendly web interface. Currently supports both Prometheus and Zeus supercomputers at Cyfronet. Allows for easy computation output sharing with group folders.

https://data.plgrid.pl/



Interested in collaboration?

We are on the lookout for collaborators interested in jointly developing new services for the life science domain. We also offer direct support in planning and carrying out scientific research. Our team takes part in collaborative, large-scale research projects as an international consortia member.

Write us at plgrid@lifescience.pl

Rimrock, one of the services of the PLGrid infrastructure, enables management of scientific computation and result handling with the use of modern interfaces based on REST (Representational State Transfer). REST is a well-established programming pattern often used in applications with distributed architectures. By using REST, access to services, applications and advanced scripts deployed on the infrastructure becomes straightforward. Main advantages of the presented service include openess to using any technology and programming language and integration with the user authorization system of the PLGrid infrastructure.

Readiness for various applications

Applying REST principles in the implementation of the rimrock service allows to use its functionalities independently of any programming language chosen to build applications on top of the computing infrastructure. It is therefore possible to create web and desktop applications as well as prepare advanced computation scripts (e.g. with the use of Bash and the curl command). An interesting approach also supported by the service is the ability to develop web applications, which can be run solely in the user web browser, minimizing the role of server-side software.

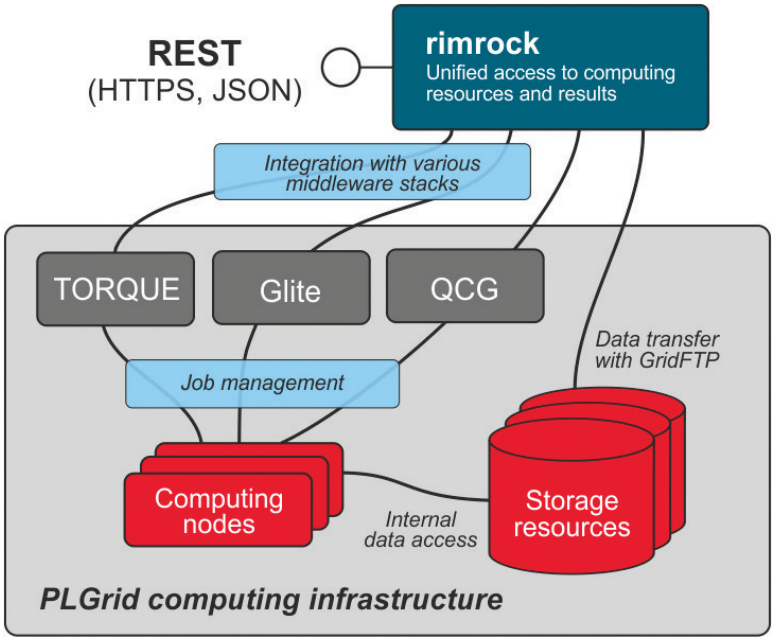
Another advantage of using REST interfaces is good support in any programming language by specialized libraries. It considerably facilitates the integration process with the computing infrastructure for both computation and data management. Each operation provided by the service is available through a single endpoint and the communication protocol uses a well-known JSON format, which also has good support in any programming platform.

The rimrock service has been successfully used during development of a web application in the domain of energy sector, allowing for harnessing the computing power of the PLGrid infrastructure for analysis of different scenarios of building a national power grid and their influence on the environment and human health. Through the integration with the rimrock service the time required for interfacing with the computing infrastructure has been minimized and the results can be obtained directly from the secure storage resources of the computing centres, without any additional data replicas.

Support for several job management systems

The rimrock service uses several job management systems, what ensures support for their unique features made available via the REST interface (e.g. using PBS directives in existing computing scripts). It allows for easy integration of legacy applications in newly developed systems. With rimrock it is possible to submit jobs directly through the PBS system (TORQUE) or with the use of one of the grid middleware stacks (gLite or QCG). A user may also submit interactive jobs by running instructions sequentially, what makes it possible to control the execution based on the returned data. Access to

computation results is facilitated by hiding the internal file transfer protocol (GridFTP) and by grouping the results according to the executed jobs.



Data security

Data exchanged with the rimrock service is transferred with secure HTTPS connections and for user authorization a temporary user certificate (so called proxy) is used. Through delegation of the mentioned certificate it is possible to constantly communicate with rimrock with just a single user login action. Furthermore, communication among several services within the PLGrid infrastructure is feasible (e.g. with the DataNet metadata management service), which allows to build complex systems on top of the existing services.

https://submit.plgrid.pl

Plgapp is a service designed for creators of web applications built on top of the PLGrid computing infrastructure. Developing applications with plgapp considerably shortens the implementation time by providing reusable elements such as handling of user login process, test and production environments for the developed application (setting up servers is not necessary any longer) and a set of libraries enabling management of computing tasks as well as access to computation results on behalf of end users.

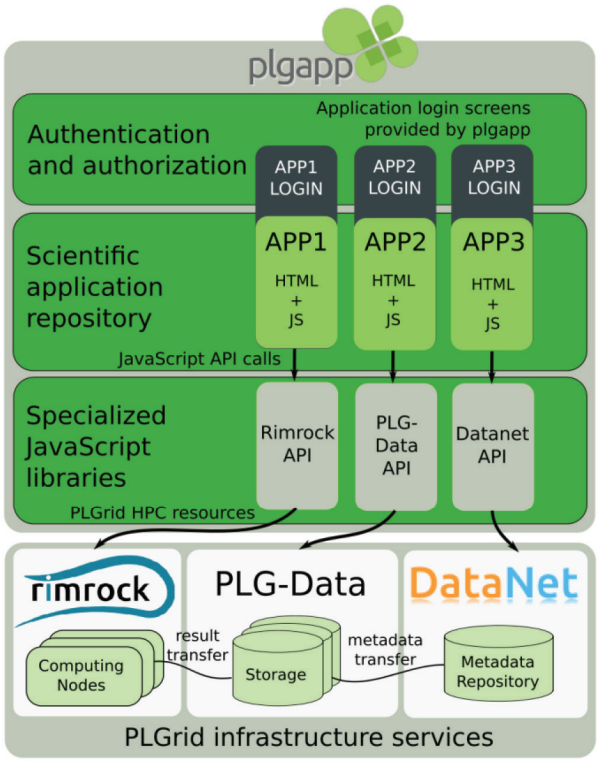
Developing web applications with plgapp

To start developing a web application with plgapp you need a text editor on your local computer and access to the service main page where you can register new web applications with a simple web form. The application right after registration is equipped with a login screen and its own address. Application code written with the text editor should consist of HTML, CSS or JavaScript files and can use any of the available Web 2.0 programming components and libraries (e.g. JQuery, Angular, HighCharts, etc.). The libraries, among other things, allow to reuse graphical components such as interactive charts, animations or even 3D visualizations on the pages of the developed application. Submitting modified application code to the plgapp server for execution can be accomplished with the help of the already mentioned web form available on the service main page or with the use of Dropbox. If a given application is implemented by several people, it is possible to create a group within which the application code can be shared. A given application at any point in time can be published to the end users with just a single mouse click on the application panel on the plgapp web page. The key aspect, which is communication with the computing infrastructure, is addressed by providing a set of specialized JavaScript libraries that can be easily added and used in the created applications.

Specialized libraries available in plgapp

Integration with the computing infrastructure is handled with a set of specialized JavaScript libraries, which allow to retrieve basic information about the currently logged in end user, submit and monitor computing jobs, send and download data files as well as manage metadata. Each library uses one of the production services of the infrastructure and therefore ensures stable and coherent programming interfaces. The presented approach enables straightforward extension of the set of specialized libraries and makes the solution open for continuously expanded service offer of the PLGrid computing infrastructure. Undoubtedly another advantage is a single location of the documentation pages for each of the specialized libraries, which greatly facilitates their use. In order to improve the learning curve of using plgapp, the documentation contains code samples for each of the libraries and web page templates to be copied and pasted.

In the picture below a structure of the applications created with plgapp is given. The parts provided by the developers of the applications are depicted with blocks titled from APP1 to APP3. As can be seen, plgapp takes care of a considerable number of application elements, which would have to be provided by the developers themselves without using the presented service.

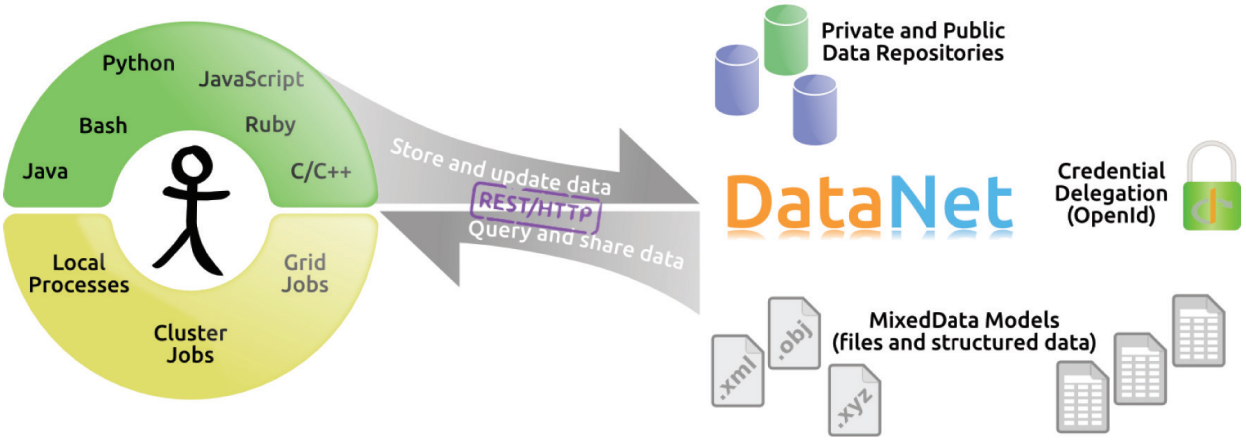


Plgapp service applications

Using the plgapp service is not limited to a particular science domain or a given scientific application. The service may be used wherever access to large computing resources is required via a modern graphical interface. As an example of using plgapp, building a portal managing simulations of reactors powered with atomic fuel can be mentioned (https://mcb.app.plgrid.pl). In this case, plgapp allowed to shorten the implementation time of the final solution and improved cooperation among developers and domain experts.

https://app.plgrid.pl

DataNet is a service built on top of the PLGrid high-performance computing infrastructure to enable lightweight metadata and data management. It allows creating data-models consisting of files and structured data to be deployed as actionable repositories within seconds. All this is accomplished with a convenient web interface where also access-restrictions to data stored in the repositories can be configured. The use of the service is not limited only to single users, but is open to others, like providers, who can access the repositories from their servers by delegating a user proxy certificate. It is even possible to gain access to the data directly from the web browser with the use of DataNet's CORS (Cross-Origin Resource Sharing) extension.



Interoperability

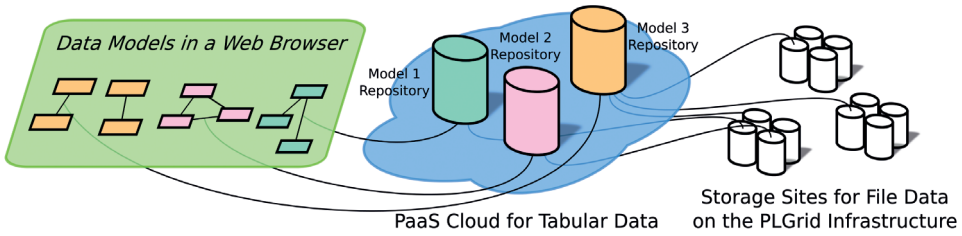
One of the main goals of DataNet is to make it usable from the largest set of languages and platforms possible. That is why we used the HTTP protocol as a basis for transferring data between computing nodes and the service, and – to make it even more convenient – we applied the REST methodology to structure the messages sent to and from the repositories. This way of transferring data has additional support in every major programming language, which makes the integration process straightforward. Our web interface has code templates for popular programming languages, which can be copied and pasted to get things going faster. DataNet is also independent of the computing location with the only requirement for it to be online, so you can use it from your local machine, a cluster job or a Grid job executed on any computing site.

Security

DataNet is fully integrated with the PLGrid authentication and authorization system, so existing users can quickly gain access to the service with a fully automated registration process. While using the service, all access rights to files and repositories are ensured and no additional privileges than those established within the PLGrid infrastructure are granted. Access to owned repositories can be configured on the level of individual users of the infrastructure. A temporary proxy certificate is used as the carrier of user credentials and it is possible to delegate it to other services for implementing complex, inter-service scenarios.

Resources

For storing metadata and data, DataNet uses resources available in the PLGrid infrastructure, thus the limits for storing capacities are specified only by individual user and group's quotas. In cases where big data volumes are processed and default user grants are not enough, a grant negotiation procedure should take place and resulting additional resources will be used by DataNet. In order to ensure user data separation, each repository is deployed on dedicated PaaS (Platform-as-a-Service), which ensures scaling and database service provisioning for structured data. For high-throughput scenarios, it is possible to configure the system to expose several instances of a given repository to increase request processing rate.



https://datanet.plgrid.pl

Scalarm is a platform dedicated for parameter studies and data farming experiments, which involve multiple executions of the same application with different input parameter values to explore its behavior in different conditions.
The main goal of Scalarm is to support such experiments in their essential stages (Fig. 1):

1. input parameter space specification,
2. application execution management on heterogeneous infrastructures,
3. results collection, analysis and visualization.

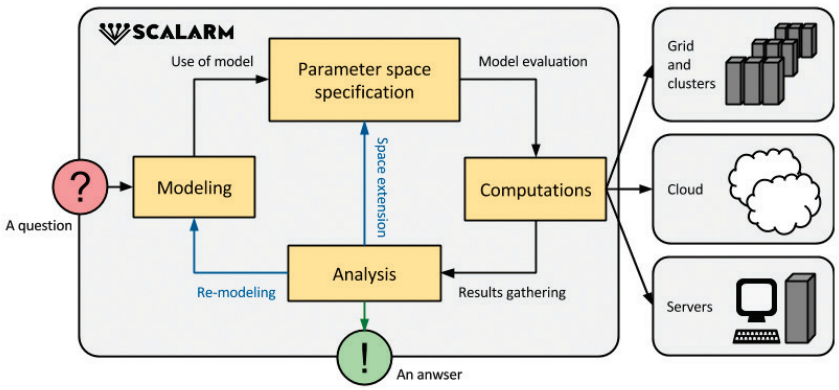


Fig 1. A diagram of experiments supported by the Scalarm platform

Input parameter space can be specified with: simple methods (single values, ranges, random values), design of experiment algorithms (Orthogonal Latin Hypercubes, 2^k), and imported CSV files. Each data point in the input parameter space constitutes a single application run in an experiment.

Scalarm can execute any scientific application including: sequential scripts, OpenMP and MPI parallel applications, and even scientific workflows by integrating with the Pegasus workflow management system. By using the adapter pattern, the user specifies how to provide input parameters, execute, monitor and collect output from the application. The results are automatically collected by Scalarm for analysis with provided graphical user interface.

Scalarm unifies access to and hides complexity of different computing infrastructures. The user decides where the application will be executed from an extensive list of supported e-infrastructures including:

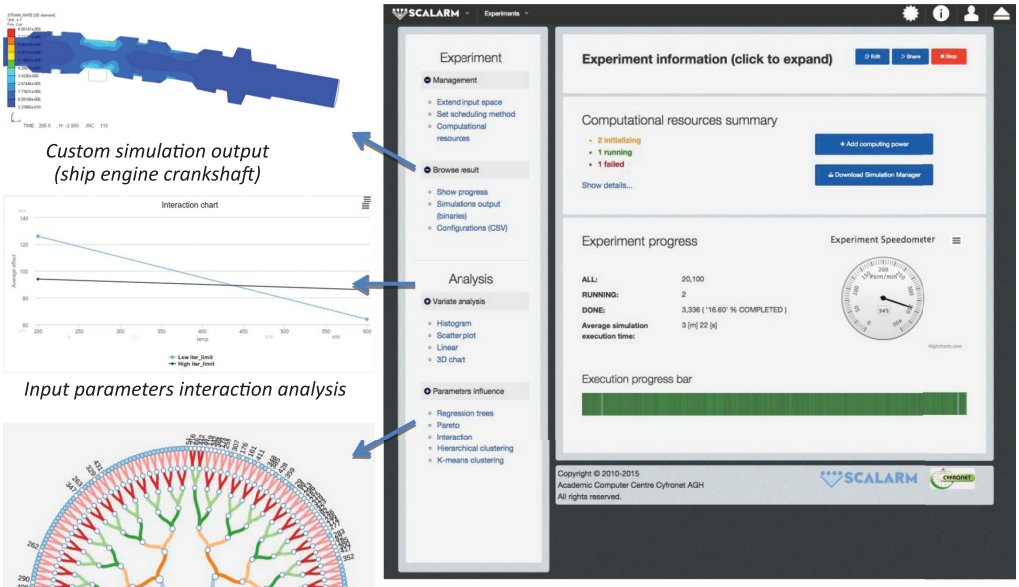


Fig 2. Scalarm web-based graphical user interface

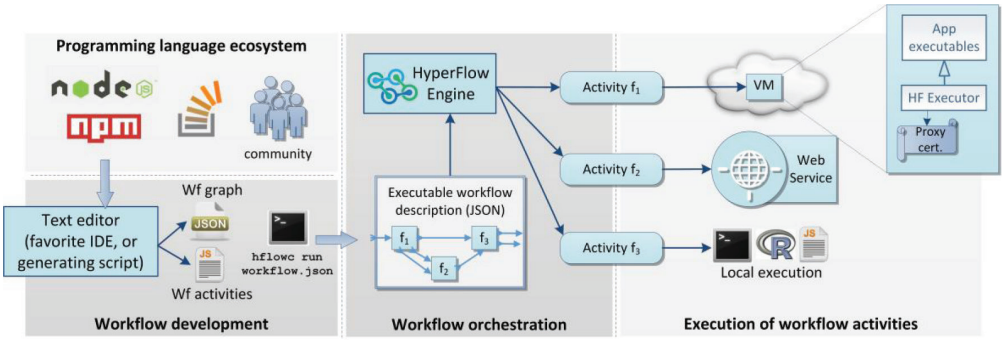
- the PLGrid infrastructure accessed via the QosCosGrid middleware,
- supercomputers with SLURM or PBS queuing systems such as Prometheus and Zeus at ACC Cyfronet AGH,
- public compute clouds – Amazon EC2, Google Compute Engine, or PaaSage environment,
- any server accessible via the SSH protocol.

Scalarm exposes a web-based graphical user interface (Fig. 2), and an HTTP-based application programming interface, which enable users to: register applications, create experiments, manage computational resources, monitor progress of the executing applications and experiments, analyze results with built-in methods. Any custom data produced by user's application can be downloaded from GUI.

Advanced features of the Scalarm platform include: a) automated input parameter space exploration with global optimization methods and sensitivity analysis techniques, b) computing resources scaling with different objective functions.

https://scalarm.plgrid.pl http://www.scalarm.com

HyperFlow is a lightweight tool that enables orchestration of scientific applications into complex pipelines or scientific workflows. HyperFlow aids users in composing their applications into workflows, deploying them in the cloud, and executing them.



Workflow programming

A workflow in HyperFlow is described as a graph of its activities (called processes) using a simple JSON-based data structure. Workflow activities perform the actual scientific procedures – steps in the scientific pipeline. In HyperFlow, workflow activities can either be implemented in JavaScript or mapped to executable programs. The JavaScript code is executed by the HyperFlow engine in the context of the `Node.js` runtime. An experienced workflow developer can thus take advantage of a mainstream programming ecosystem – large community, advanced tools, thousands of libraries and other resources – instead of using a proprietary development environment. Consequently, workflow activities can easily be programmed to invoke external Web Services, or execute local commands as part of the scientific pipeline defined by the workflow.

In the second option, the workflow developer can choose not to implement any JavaScript code, only associate each workflow activity with a previously prepared Virtual Machine image where appropriate programs are installed, and specify commands that are to be executed when a given workflow activity is triggered.

The availability of these two programming approaches makes HyperFlow equally suitable for experienced programmers / software engineers who desire low-level programming capabilities and high productivity, and domain scientists who are not experts in IT technologies and only wish to construct scientific pipelines out of existing modules.

Workflow deployment

HyperFlow automates workflow deployment in the cloud. The user only needs to prepare a configuration file specifying the mapping of workflow activities onto available Virtual Machine images, while the HyperFlow tool takes care of the rest. The user invokes a simple command hflowc setup which results in creation of appropriate VM instances in the cloud. These VM instances contain the workflow runtime environment and the scientific applications invoked from the workflow.

Workflow execution

After the workflow instance has been created in the cloud, the user executes the workflow simply by invoking hflowc run <workflow_directory>. Every workflow runs with its own instance of the HyperFlow runtime environment. Consequently, different workflow runs are isolated from each other which increases security and reliability.

The HyperFlow cloud runtime environment (called HyperFlow Executor) automatically takes care of transferring input data from the user directory to Virtual Machine instances, invokes the application executables and uploads output data back to the user directory. A variety of data transfer options are available, including a network file system, secure gridftp, and Amazon S3.

Applications

HyperFlow has become a part of several larger systems where it has been used for a number of applications. In the PLGrid infrastructure, HyperFlow serves as a workflow management system that enables the users to run scientific workflows in the cloud. An example application is a workflow-based solver for finite element meshes which can be applied to diverse problems. HyperFlow is also being integrated with the PaaSage middleware (http://www.paasage.eu) as an execution engine for scientific applications deployed in a multi-cloud environment. In the ISMOP project (http://www.ismop.edu.pl), HyperFlow is a component of a flood decision support system used to orchestrate flood threat assessment workflows.

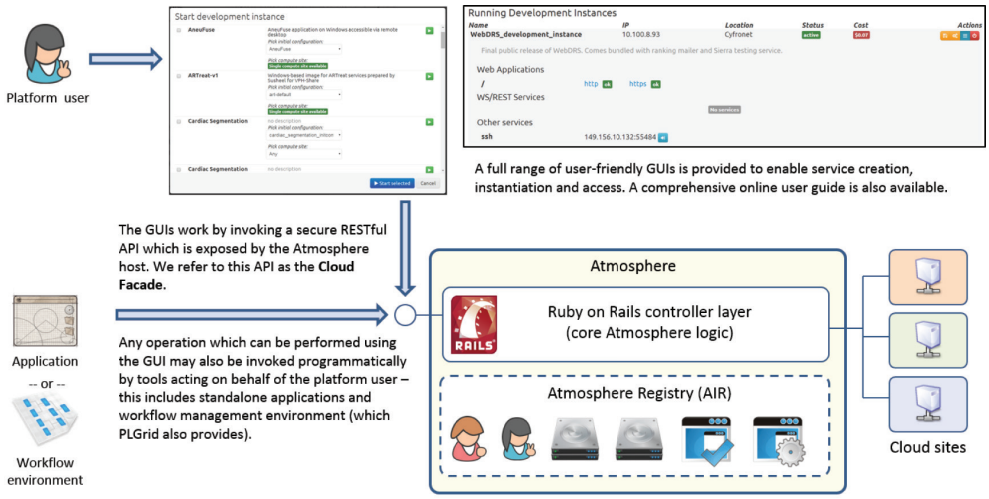
Contact

HyperFlow is developed and maintained by the DICE team (http://dice.cyfronet.pl). Please feel free to contact us in case of any questions or suggestions.

HyperFlow source code and manuals are available at https://github.com/dice-cyfronet/hyperflow.

Atmosphere

The Atmosphere Cloud Platform is a one-stop site for management and interaction with the computational cloud resources operated by the PLGrid e-infrastructure. Atmosphere is a user-friendly environment where hybrid cloud resources contributed by various participating institutions and sites (including public clouds) are seamlessly integrated into a coherent, unified resource space, made available to PLGrid users. Accessing the Atmosphere platform enables you to exploit the cloud computational resources which are part of PLGrid.



Cloud service abstraction

The principal goal of Atmosphere is to make interaction with cloud sites easy for beginners and experienced users alike. Atmosphere can deploy virtual machines into the cloud, create image snapshots and support sharing of computational services by PLGrid teams without the need to use any low-level cloud service libraries. A convenient GUI encapsulates all features offered by Atmosphere to each class of users: system administrators, application developers and end users. In addition, a set of APIs is provided to enable integration of the Atmosphere platform with external applications, tools and workflow management systems.

Security

The Atmosphere cloud platform is integrated with PLGrid authentication and authorization mechanisms. All users of the PLGrid infrastructure can request access to cloud resources simply by joining the plgg-cloud2 user group. Thereafter, each user can create, use and share cloud-based virtual machines in the context of their own research team, or for individual research purposes. The visibility of each virtual machine (and the corresponding VM images) is restricted to the PLGrid team, in which the machine was created. Atmosphere performs automatic billing and resource consumption auditing for all VMs.

Resources

Atmosphere can interact with many different types of cloud resources contributed by individual cloud platforms. For the purpose of PLGrid, a dedicated cloud site has been set up at ACC Cyfronet AGH, comprising of computational nodes managed by the OpenStack cloud middleware. Atmosphere can also interact with public cloud providers, such as Amazon, RackSpace, Google Compute and many others. All this is done without forcing the user to learn any technology-specific libraries or APIs.

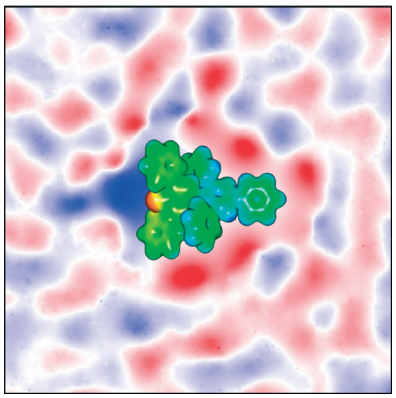
Applications

In addition to raw OS templates for service developers, the PLGrid cloud site supports a variety of ready-to-use applications – such as the Hyperflow workflow services for the ISMOP project, capable of farming out urgent computational tasks in the context of flood embankment monitoring. Any web or REST service can be deployed into the cloud using Atmosphere. It is worth noting that an earlier version of Atmosphere, developed in the framework of the VPH-Share project (http://vph.cyfronet.pl/tutorial) was successfully exploited in the Virtual Physiological Human community by approximately 25 research teams affiliated with the VPH-Share and VPH-DARE consortia, as well as external partners who collaborate with ACC Cyfronet AGH. Additionally, Atmosphere is now used to provision computational services for the EurValve project in which ACC Cyfronet AGH is a member.

https://cloud.plgrid.pl

Chemistry and Biology – electronic structure and molecular dynamics software

Modern computational chemistry requires constantly increasing resources. More and more computational power is needed to make large systems (especially those being current challenges of nanotechnology or biological sciences) tractable and improve the accuracy of obtained results. Fortunately, constant progress in computer technology together with specialised software offered by Cyfronet meet this demand and enables various kinds of chemical computations.



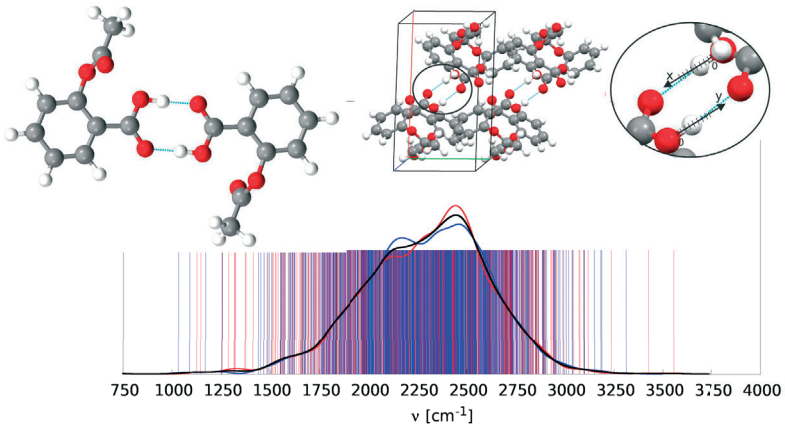
A. Eilmes, P. Kubisiak: Electrostatic potential of an ionic liquid around the solvated dye molecule

Zeus and Prometheus' nodes provide up to 256 GB of RAM and 64 cores per physical node, which enables quantum chemical computations that require large amount of memory or high number of cores with shared memory. Moreover, fast InfiniBand interface allows good speed-up of calculations if distributed over many nodes. Various quantum chemistry codes also need fast and broad I/O to storage systems. The parallel-distributed Lustre scratch file system and possibility to use RAMDisk on selected nodes enable that.

Efficient quantum chemistry computations rely also on efficient installation of scientific software and its proper usage. Our administrators' team has got necessary skills, knowledge and vast experience in installing various applications and running computations efficiently. Our board portfolio of software used in chemistry contains many packages. Among them there are:

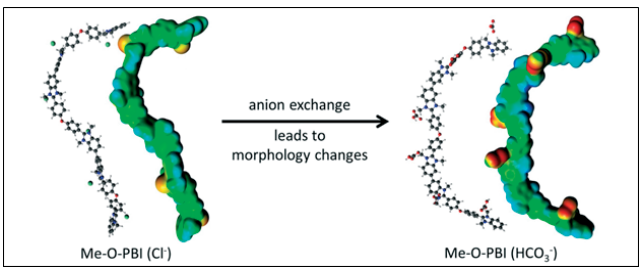
- Versatile and widespread used quantum chemistry codes such as Gaussian, GAMESS UK, NWChem, Jaguar, Q-Chem and TURBOMOLE, which are capable of calculating electronic structure and various properties of diverse molecular systems using both ab initio, density functional theory and semi-empirical methods.
- Molpro, CFOUR and Dalton suites to analyse chemical systems with great accuracy using sophisticated methods such as CC (up to CCSD(T)) and MCSCF.
- ADF molecular modelling suite, which provides methods to examine various properties (especially spectroscopic, such as NMR and ESR spectra) of molecular systems with reliable relativistic ZORA approach, COSMO-RS method and all-electron basis sets for the whole periodic table. With addition of versatile and well-constructed GUI (ADFInput, ADFViev, etc.) ADF package is used by many of our users.
- Several packages could be used for solid-state systems. Among them BAND, Quantum ESPRESSO and SIESTA are worth mentioning.

- Desmond, Gromacs, Amber, LAMMPS, NAMD, CPMD, CP2K and Terachem suites for molecular mechanics and molecular dynamics simulations of systems containing hundreds of thousands and more atoms.



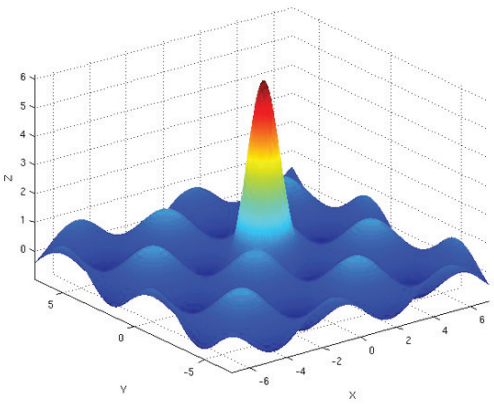
Description of proton dynamics of strong hydrogen bonds in aspirin crystals. Published by M. Brela, M. Wójcik, Ł. Witek, M. Boczar, E. Wrona, R. Hashim and Y. Ozak in J. Phys. Chem. B 2016, 120, 3854–3862

Nowadays general-purpose computing on graphics processing units (GPGPUs) in many scientific domains provides great speed-up of calculations (up to several orders of magnitude). In our computing Centre some of nodes provide possibility of such calculations on CUDA enabled GPGPUs. Among software prepared to run on graphical processors our administrators' team prepared quantum chemical packages such as GAMESS, Terachem, NAMD, Quantum ESPRESSO and Adorycja. Computational chemists from Faculty of Chemistry of Jagiellonian University and our experts working in our supercomputing Centre develop the last one.



Electrostatic potential of molecules in anion exchange membrane. Published by W. Germer, J. Leppin, C. Kirchner, H. Cho, H. Kim, D. Henkensmeier, K. Lee, M. Brela, A. Michalak and A. Dyck in Macromol. Mater. Eng. 2015, 300, 497–509

CAD/CAE applications



Computer-Aided Design and Computer-Aided Engineering applications are nowadays essential tools in the process of developing and building almost everything – from car parts to buildings. Through computer simulations, engineers can check durability of constructs and devices; perform linear and non-linear structural analyses of contact phenomena, plasticity, recoil, etc. CAD/CAE software provides analysis of thermal conductivity, radiation and phase shifts. Significant for science are also fluids simulations: velocity fields, pressure fields, heat distribution, chemical reactions, etc.

Cyfronet's users can resolve all these tasks thanks to CAD/CAE packages of ANSYS, ABAQUS, FLUENT, MARC and OPERA.

ANSYS is a complex structural simulations package with intuitive graphical user interface, supporting scientists from nearly any area of science or business. Results are calculated with high precision and may be presented by plots or tables, for example isosurface diagrams and deformations. Computational capabilities of ANSYS are very high and involve: harmonic and spectral analysis, statistics and dynamics.

ABAQUS is devoted to solving problems in industry using finite-elements analysis. A user can prepare a combination of finite-elements, materials, procedures of analysis and sequences of loads, according to individual requirements, to simulate vehicle loads, dynamic vibrations, multibody systems, impacts, crashes and much more.

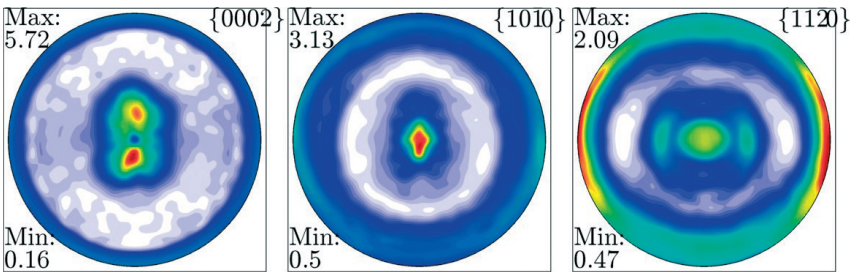
FLUENT software offers the broad physical modeling capabilities needed to model flow, turbulence, heat transfer and reactions for industrial applications ranging from air or liquid flow to semiconductor manufacturing. FLUENT can be used in numerous science domains, including chemistry, metallurgy, biomedicine, electronics, material design and many others.

MARC is a general-purpose, non-linear finite element analysis solution to accurately simulate the product behavior under static, dynamic and multi-physics loading scenarios. It has capabilities to simulate all kinds of non-linearities, namely geometric, material and boundary condition non-linearity, including contact. It is also the solution that has robust manufacturing simulation and product testing simulation capabilities, with the ability to predict damage, failure and crack propagation. All that can be combined with its multi-physics capabilities that helps couple thermal, electrical, magnetic and structural analyses.

OPERA is a finite element software suite for design and optimization of electromagnetic devices in 2D/3D. It gives accurate numerical solutions for problems from multiple areas of science, including electrostatics, magnetostatics, low and high frequency electromagnetics. The software gives an ability to design and optimize many types of electrical devices: transformers, motors, switches, micro-machines, MRI scanners and X-ray tubes. It is a powerful virtual prototyping facility to accelerate the design process.

Symbolic math applications

Mathematical applications enable to conduct in reasonable amount of time even very complex and complicated calculations. Users of ACC Cyfronet AGH have access to software that supports calculations in the field of algebra, analysis, combinatorial math, statistics, theory of numbers, geometry or other math areas. Running calculations like integration, differentiation, symbolic processing, matrix operations, approximation and interpolation, Fourier and Laplace Transforms, digital signal processing, etc. is a lot easier. Results can be visualized with appropriate tools. Some of the applications can create interactive 2D and 3D plots. In scientific work, preparation of precise model that most accurately describes analysed issues, is essential.

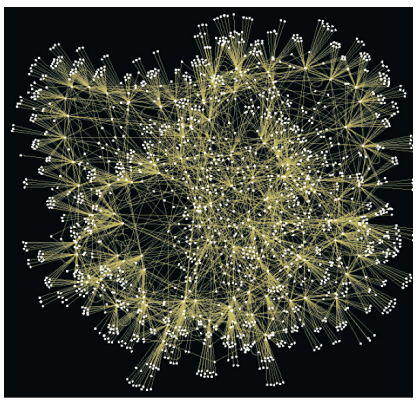


Bartosz Sułkowski: Texture of rolled AZ61 magnesium alloy. Pole figures calculated by Mtex, the Matlab package.

A good example of software environment, which can be applied in above-mentioned issues, is MATLAB. Its modules (Toolboxes) allow performing computations in the field of financial modelling, partial differential equations, linear and non-linear optimization and much more. It is also possible to use Simulink – the environment oriented for simulations and visualizations from blocks, without the need for traditional programming.

Apart of that environment, users can find in our software a useful application, MATHEMATICA, which allows parallel computations with defined precision, dedicated for symbolical and numerical calculations. An advantage of MATHEMATICA is, among other things, a tool for fixing mistakes.

Another example of universal and interactive mathematical software is MAPLE. It can be used for simplification of expressions and symbolic processing. It offers databases, enables code generation in other programming languages, creating slideshows with user commands and communication with MATLAB and CAD systems.



Rafał Rak: One minute price returns network for KGHM (the Polish stock company)

HPC Users' Conference (KU KDM)

The basic premise of the HPC Users' Conference was to initiate annual scientific meetings devoted to users performing computations in ACC Cyfronet AGH with use of high-performance computers, computing clusters and installed software.

The conference was launched in 2008 and included several presentations by Cyfronet employees – describing the resources available in the Centre, as well as numerous lectures of researchers – presenting the scientific results achieved using Cyfronet hardware and software. In addition, two invited speeches were given by Norbert Attig from Jülich Supercomputing Centre and Jaap A. Kaandorp from University of Amsterdam.

The first edition of the conference attracted much attention and increased the interest of users in Cyfronet resources. It proved that this type of event was much awaited and needed.

Nowadays, the HPC Users' Conference focuses on the large-scale computations and simulations, novel algorithms in computer science, tools and techniques relevant to high-performance computing, teaching in computer science, databases. However, the main aim of the conference is the overview of research results carried out using the computer resources of Cyfronet. It is also an opportunity to familiarize the users with the Centre and its resources, including the PLGrid infrastructure.

The conference includes a series of talks by scientists who perform research using Cyfronet resources and can present the role of these resources, typical usage scenarios and performance aspects. The event is an important opportunity for Cyfronet representatives to meet with these scientists and acquire the knowledge necessary to take the proper actions in order to adapt the computing infrastructure to scientists' needs and fulfil their requirements. On the other hand, the conference also gives a possibility for researchers representing various disciplines to exchange experience and

become familiar with the new technologies and domain-specific services currently being deployed at the Centre.

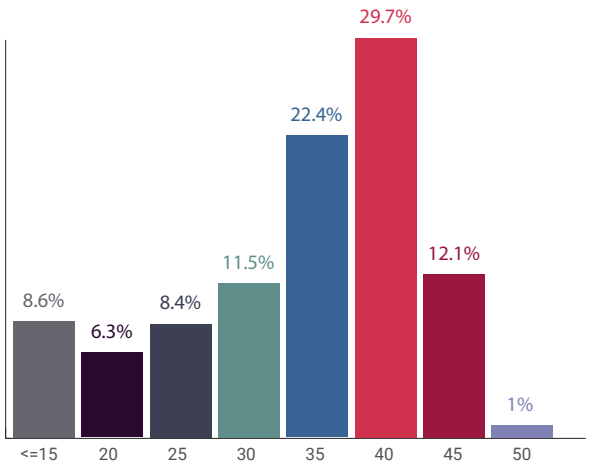
The crucial parts of the conference are meetings with suppliers of Cyfronet hardware and software, as well as the panel discussion on efficient use of these resources. The latter is always attended by users – researchers, who use the chance to get familiar with news regarding the computing infrastructure in the Centre and to inform Cyfronet experts about issues encountered while interacting with this infrastructure.

The conference is accompanied by poster and training sessions – in 2018 the training was focused on the machine learning using the Python language.

Contributed papers elaborated on the basis of the best conference talks were published in one of two well-regarded IT journals: Computing and Informatics (CAI) (http://www.cai.sk) or Computer Science (CSCI) (www.csci.agh.edu.pl).

http://www.cyfronet.pl/kdm18/

Ministry of Science and Higher Education
marks of articles published in 2017
by Cyfronet Users in scientific journals



CGW Workshop

The CGW Workshop was organized for the first time in 2001 by ACC Cyfronet AGH, in cooperation with the Institute of Nuclear Physics Polish Academy of Sciences and the School of Banking and Management in Krakow. This event was related to the Cyfronet's participation in a large, international project CrossGrid (2002-2005), funded by the EU. Among the guests invited to the conference were eminent scientists, experts in the field of grid technologies: Fabrizio Gagliardi, Wolfgang Gentzsch, Jarosław Nabrzyski and Piotr Bała.

Since then the conference has been organized annually and has become a very important event in Europe. Currently, the organizer of the conference is ACC Cyfronet AGH and Department of Computer Science AGH.

CGW Workshop is an opportunity for the presentation of research and development activities, supported by large-scale simulations, using grid and cloud technologies, and other computing techniques. The conference also provides the overview of research carried out in the EU and national projects, addressing distributed infrastructures.

In particular, scientific and technical achievements within the PLGrid Programme are presented, together with those related to the use of the PLGrid infrastructure.

The conference program includes invited lectures and oral presentations of participants. In parallel, the poster session is organized, as well as the exhibition stand with the materials promoting the EU grid projects, and the ones realized under the PLGrid Programme. The conference is also

accompanied by training aimed at familiarizing the participants with the latest tools and platforms, which facilitate the use of grid and cloud infrastructures.

Contributed papers accepted for presentation at the conference provide a very good overview of the research activity in the area of e-Science and distributed computing infrastructures.

In addition to publication in the CGW Workshop proceedings, the extended versions of selected papers are printed in the Computer Science Journal (CSCI), published by the AGH University of Science and Technology in Krakow. The journal is prepared in cooperation with many renowned computer science researchers from all over the world and its main purpose is to create a forum for exchanging research experience for scientists specialized in different fields of IT.

The Steering Committee of the CGW Workshop makes every effort to address during the conference such topics, which are in line with the latest trends in IT. The result of these endeavours is the fact that each year the conference has a growing interest of researchers involved in the creation and development of information technologies, as well as the users of these technologies.

http://www.cyfronet.pl/cgw18/

EOSC-Hub – a milestone in the European Open Science Cloud implementation

European Open Science Cloud (EOSC) is an initiative that aims to federate existing European e-infrastructures and research infrastructures, in order to provide scientists with a complex environment for storing, processing, analysing, finding and sharing data. More specifically, it will give researchers an access to all publically funded research data in Europe, across disciplines and borders.

The role of EOSC-Hub

EOSC-Hub is to be a heart of the future EOSC, serving as its integration and managing system. In the first instance it aims to integrate and consolidate resources of EGI Federation, EUDAT CDI, INDIGO-Data Cloud, and then other major research e-infrastructures. The project's expected outcome is a comprehensive catalogue of services supporting advanced data-driven science (including platforms, policies and managing systems), with a pan-European, unified access mechanism via secure Authentication and Authorisation Infrastructure (AAI).

One of the EOSC-Hub key tools is Marketplace – a user-facing platform where services of different nature: infrastructural, application, human-based can be promoted, discovered, ordered and accessed. A set of functionalities implemented in the Marketplace supports efficient service and order management and facilitates the interactions of a user with a service provider. As a result, researchers and other users who can gain from EOSC-hub services have a single platform giving them opportunity to find means supporting their research and gain access to them. The technological solution behind the Marketplace was delivered as a result of EGI Engage project by ACC Cyfronet AGH, who was also responsible for designing and creating a prototype of Open Data Platform. Cyfronet's contributions will help to integrate different data repositories available through the EGI platform and link them to OpenAIRE infrastructure.

Other part of EOSC-Hub, Operations Portal, supports all EGI Operations Centres and provides such capabilities as infrastructure oversight, security operations, broadcast or availability reporting. It's also designed to meet the needs of Virtual Organizations' management. EOSC-Hub also contains systems monitoring the software versions installed at data centers participating in the project. By ensuring that the entire infrastructure runs on up-to-date software, it increases the security and efficiency.

By integrating different e-Infrastructures' resources and making them more visible and accessible, EOSC-Hub will support scientific research and development of the multi- and interdisciplinary collaboration across the borders.

From the project launch at the turn of 2017 and 2018, EOSC-Hub has obtained 100 partners from 53 countries. At the moment it offers nearly 50 services in 13 work packages. EOSC-Hub will run until 2020 with funding from the European Union's Horizon 2020 research and innovation programme, as well as from contribution of participating units.

A constant development of EOSC-Hub and EOSC is supported by European supercomputers (e.g. Prometheus) united within EuroHPC agreement.

European High-Performance Computing – EuroHPC

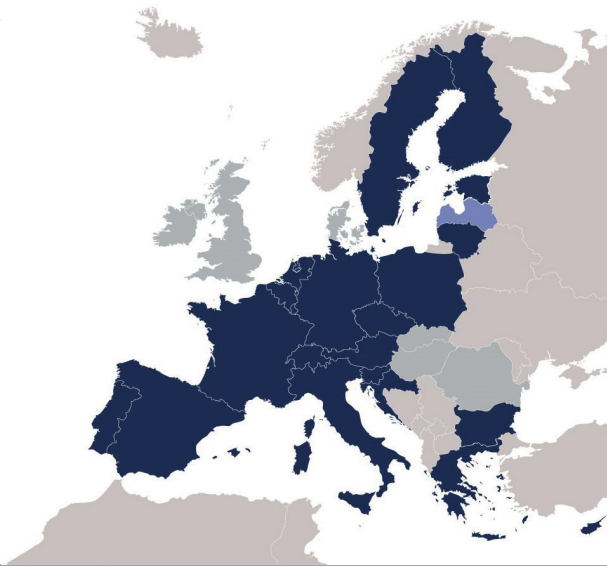
The idea standing behind the EuroHPC is to build and deploy at first instance two pre-exascale supercomputing machines (until 2022) which will be followed by building the true exascale (1-2 EFLOPS) supercomputer by 2025. The exascale means that the machine will be able to perform 10^{18} floating point operations per second. In comparison, the existing petascale (PFLOPS) machines are two/three orders of magnitude lower in performance (10^{15} calculations per second).

Until the end of September 2018, 24 European countries, including Poland, have committed to participate in the EuroHPC Joint Undertaking – an agreement was formally established at the end of September 2018, with the headquarters in Luxembourg. In addition, EuroHPC welcomes also private members, like representatives from HPC and Big Data stakeholders from both academia and industry.

The EuroHPC initiative aims to support two main pillars related to exascale computing. One focussing on activities for the acquisition, deployment, interconnection, operation and access time management of world-class supercomputing and data infrastructures, while the second one is related to the Research and Innovation (R&I). These, include activities being a key to establish an innovation ecosystem addressing hardware and software supercomputing technologies and their integration into exascale supercomputing systems, advanced applications, services and tools, skills and know-how.

The total cost of the initiative will reach almost 1 billion EUR, out of which around half will be covered by the EU contribution and the similar amount by Member States, associated countries and private investors.

Together with representatives of Ministry of Science and Higher Education, ACC Cyfronet AGH is coordinating activities related to design and implementation of the EuroHPC Joint Undertaking.



Signatory European Countries with Latvia, which has also committed to participate

The aim of the EOSC-hub project is to prepare the launch of a production infrastructure for open science in Europe and the practical application of solutions developed as part of the EOSC-Pilot project to a real large-scale environment scattered across most European countries.

The solutions developed within the project will be a breakthrough step in the creation of innovative, exascale data processing services, maximizing the benefits of modern data processing systems.

The goal of the project is to build specialized solutions for managing and processing large-scale data in a hybrid cloud, thus introducing access and data migration in distributed cloud environments.

The project will develop a Business Case to establish in Poland Europe's first Centre for simulation-driven healthcare Decision Support. The Centre's mission comprises development of new computation-based solutions for diagnostics and therapy in daily healthcare to improve personalised patient treatment, enhanced growth of SMEs which deal with technologies and services for novel personalised diagnostics and treatment strategies, as well as strong advancement of algorithms, models and technologies involved in personalised medicine.

The overall objective of SERA is to give a significant contribution to improve the access to data, services and research infrastructures, and deliver solutions based on innovative R&D in seismology and earthquake engineering, aiming at reduction of the exposure of our society to the risk posed by natural and anthropogenic earthquakes.

The goal of AARC2 is to design an AAI (authentication and authorisation infrastructure) framework to develop interoperable AAI, to enable researchers to access the whole research and infrastructure service portfolio with one login.

The objectives of PRACE-5IP are to build on and seamlessly continue the successes of PRACE and start new innovative and collaborative activities proposed by the consortium.

The EPOS Implementation Phase project (EPOS IP) builds on the achievements of the successful EPOS preparatory phase project (EPOS PP). The EPOS project is integrating the diverse, but advanced European Research Infrastructures for solid Earth Science, and will build on new e-science opportunities to monitor and understand the dynamic and complex solid-Earth System.

The aim of the project is to develop a comprehensive, clinically-compliant decision-support system to meet the challenge of treatment optimisation in case of the Valvular Heart Disease, by quantifying individualised disease severity and patient impairment, predicting disease progression, ranking the effectiveness of alternative candidate procedures, and optimising the patient-specific intervention plan.

The aim of the EGI-Engage project was to accelerate the implementation of the Open Science Commons by expanding the capabilities of a European backbone of federated services for computing, storage, data, communication, knowledge and expertise, complementing community-specific capabilities.

The aim of the project was the development of a model-based predictor system, supporting the flexible design of strip rolling, joining functionality of numerical simulations, material modelling, sensitivity analysis and optimization.

The main objective of the VPH-Share project was creation and implementation of a platform for the sharing of scientific applications implemented by the members of the European Virtual Physiological Human (VPH) consortium, as well as to adapt existing applications and data repositories for sharing within the framework of the aforementioned platform.

The project aim was to develop and implement new tools and services used to run interactive applications, which required high computing power and large data collections in the grid environment.

The main objective of the project is to provide satellite data coming from the Sentinel satellites of the Copernicus network. The project will create an infrastructure for automatically downloading data directly from satellites, their secure storage and sharing for the purposes of science, administration and training.

Within the project it is planned to develop a comprehensive diagnostic test based on analysis of the material obtained from the tumor via surgery or biopsy. This will lead to a liquid biopsy-type method based on tumor cfDNA in blood, which could be a quick and minimally invasive alternative to biopsy.

The project aims at building the national research infrastructure for solid Earth Science and its integration with international databases and services implemented under the European Plate Observing System (EPOS).

The objective of the project was the development of the specialized technological competence centre in the field of distributed computing infrastructures, with particular emphasis on grid technologies, cloud computing and the infrastructures supporting calculations on large data sets. As a result, a great computing power and huge storage for digital data were offered to users. They also obtained access to a set of basic and end-user services, allowing for easier integration of their solutions, specific to the selected fields of science, with the PLGrid infrastructure.

The project aimed at integration of new groups of researchers from the next 14 disciplines with the PLGrid computing infrastructure. Thanks to deployment of new domain grids, research teams from these areas are able to faster obtain results of their calculations and to better integrate with the national computing infrastructure for science.

The project supported, by means of IT, Polish research teams in conducting research and also enabled extensive collaboration among these teams, as well as international cooperation in the area of e-Science.

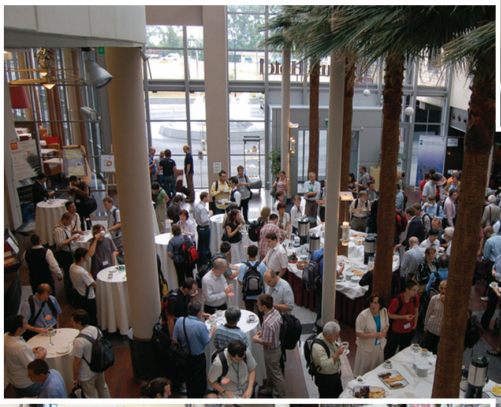
The preparatory phase of Polish input for Cherenkow Telescope Array Project. This project has been undertaken to help design high level tools easing CTA data analysis. The main area of work covered the development of InSilicoLab Science Gateway for CTA with focus on job results analysis and visualisation, in particular.

The aim of the project was integration of selected services available in the PIONIER network, and the development of the new services, e.g., with the increased reliability and security.

The project envisaged the creation and launch of five services running on the basis of the PIONIER network. These services included: video conferencing services, eduroam services, campus services, universal archiving services and scientific interactive HDTV services.

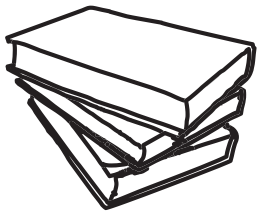
The aim of the project was the development of 21 environmental science data communication networks providing the scientific institutions across the country with access to a modern and secure network infrastructure, supporting the research and development of Polish groups of scientists.

Within the project the Polish Grid Infrastructure (NGI) has been built to provide the Polish scientific community with an IT platform based on computer clusters, enabling research in various domains of e-Science.




www.CYFRONET.krak

Work of young scientists in Cyfronet



The contest for the best PhD thesis conducted with the help of computing resources of ACC Cyfronet AGH is a tradition in our Centre. Organized annually, it is becoming an important event promoting research performed by young scientists. For the 2018 edition of the Contest, participants submitted many PhD theses focused on variety of scientific problems in mathematics, material engineering and robotics, physics, chemistry and others. Also the utilisation of the resources varies, as the contesters used different parts of the Cyfronet's computing infrastructure: both supercomputers as well as clusters and the GPGPU platform.

The laureates of the Contest are invited to give a talk during Cyfronet's Open Day. We are honoured to present here selected interviews with the Contest participants.

Join the 2019 Contest edition!

http://www.cyfronet.pl/konkurs

The laureate of one of the previous Contest editions

Maciej Bendkowski

The interview with the author of the PhD thesis: "Quantitative aspects and generation of random lambda and combinatory logic terms".

Why in your opinion it's important not only to learn mathematics, but also to develop the discipline?

Mathematics is the language of modern science, be it engineering, natural sciences, medicine, economics, or even humanities. It gives its users the ability to organise thought, formulate rigorous arguments, and distil the essence of considered problems. Consequently, it is absolutely central to humanity's endeavour of understanding our surrounding reality. Without advances in fields like theoretical computer science, number theory or combinatorics we would live in an utterly different world, without most of the recent technological benefits we rely on everyday such as, for instance, the Internet. Mathematics is not only the language of science, it is the future of science.

What influenced your decision to dedicate your PhD thesis to combinatory logic and related terms? Why did you find it challenging?

During my student years, I grew fond of formal logic and the theory of computability, taught by Prof. Marek Zaionc at the Jagiellonian University. Related methods of automated, formal reasoning and questions regarding decidability of certain decision problems struck a chord with me and so I decided to pursue this direction under Prof. Marek Zaionc, who later become my supervisor.

What were the biggest obstacles and how did you overcome them?

Due to the atypical combinatorial structure of some of the investigated term properties, such as normalisation or typeability, none of the classical techniques applied and some alternative ones had to be developed. As an example, consider the following problem. Given a program, what are the chances that it will eventually terminate and finish its computations? Is it true that almost all programs terminate and just a small, insignificant proportion of them continues their computations forever?

Together with my co-authors, we had to find an alternative approach to the posed problem and, in effect, approximate undecidable properties with recursive sub- and supersets. Our task appeared intrinsically difficult. Fortunately, one day I experienced an eureka moment where I suddenly realised that it was possible to investigate programs terminating in a fixed, bounded amount of time. Moreover, I discovered that it was possible to automatise my technique and turn it in a program itself. Consequently, I was able to give an effective approximation scheme and explain not only the structure of terminating combinatory logic terms, but also explain the reason for the intrinsic difficulty of the initial problem.

How did ACC Cyfronet AGH resources help with your calculations?

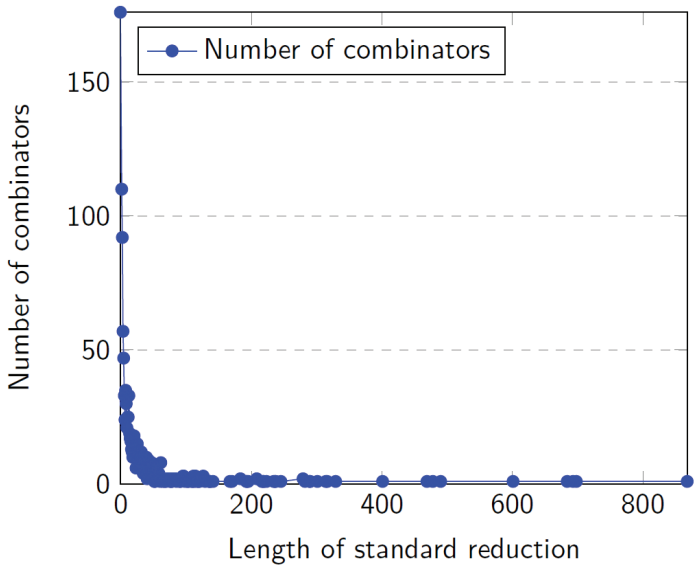
Although using my method it is possible to automatically refine the estimate for the likelihood that

a program expressed in the language of combinatory logic terminates, due to the sheer amount of computational resources involved in the calculations, obtained lower and upper bounds are still quite far away. In order to evaluate the developed approximation scheme, I needed to find an alternative method of estimating the ground truth. For that purpose, I decided to resort to Monte Carlo techniques and use the Prometheus supercomputer in my computations.

In principle, the experiment was quite simple. I generated large, uniformly random (conditioned on size) combinatory logic terms and normalised them for an a priori fixed number of steps. Moreover, I assumed that if the computations took too long, the represented computation would never terminate. Of course, this assumption cannot be always fulfilled, however since the problem at hand is undecidable, it was justifiable for my experimental purposes. The Monte Carlo scheme was implemented in Haskell, and parallelised over several dozens of available CPU cores. Without this massive parallelisation, required computations would take a few months. Fortunately, with Prometheus it was possible to finish them in a couple of days. The Monte Carlo experiments suggested that there exists a large gap between the currently best known theoretical bound and the conjectured truth.

What outcome may your research have in the future?

In recent years, we have seen a growing interest in using randomness in the process of software testing. In this approach programmers do not write tests themselves, but rather specify certain invariants and properties which their code is assumed to satisfy. It is the computer's task to generate random test cases and try to falsify the programmer's claims. It has been shown that such a testing method is quite effective and provides strong evidence that tested code obeys declared invariants. Together with my co-authors we discovered that the program representation strongly influences the properties of typical programs. It is my hope that this research will eventually be used in advancing software testing tools and help programmers in their daily work writing correct and safe code.



Random distribution of the length of random reductions of SK-combinators. Data obtained with the use of Prometheus supercomputer

What could you advise younger colleagues, who are at the beginning of their PhD studies?

Modern science is more international and cooperative than ever before. Internet has greatly mitigated the collaboration process; international borders and physical distance are no longer a truly prohibiting force. I would encourage younger colleagues to actively seek new opportunities and contact other scientists outside their local circle. Science is a global scale endeavour and, as such, requires global scale efforts.

Krzysztof Grabowski

The interview with the author of the PhD thesis:
„Design and development of the sensors for structural health monitoring (SHM) based on the carbon nanomaterials".

How did you get interested in new materials?

I must say that it happened thanks to my supervisor, who started to guide me already during work on my MSc thesis. Then, the idea appeared to create an intelligent surface structure, which would be able to measure strains, and to armour machines with such a skin, thanks to which they could also „feel".

Professor Uhl wanted to use graphene for this purpose, but while I was initially analysing potentially useful materials, I came across carbon nanotubes.

While talking of surfaces, what kinds of them do you mean?

We're talking about mechanical structures, especially such as ships' hulls and planes' plating. There were some queries from the industry, like from a yacht manufacturer from Germany, who seeks for a way to monitor ships during transport, in order to check, whether the yacht has been hit or not. There were also some ideas to monitor strains of bridges.

How would you make a non-specialist more familiar with your research?

The whole idea is that nanotubes are sunk in epoxy resin. This way we get a kind of a net, which we strain. When we do that, the resistance changes. Of course, there are many other factors, but I try to keep it simple for better understanding. Creating such structures in an experimental way is in fact easy. But it's more complicated to ascertain the real strain after the measured resistance change.

So, from the experiments we enter the stage of calculations. How were the Cyfronet's resources helpful with performing them?

The sensors represent the macro-scale. I was interested in dimensions, let's say, 10x10mm. It's very difficult to estimate the number of carbon nanotubes with the diameter of 6 to 9 nanometres on a surface of this size. Performing all calculations without Prometheus would be virtually possible, but as a result my PhD project would have been finished after 5 more years, of course, only on the assumption, that all data was entered correctly and I wouldn't have to start calculations once again.

What other difficulties have you encountered?

In the experiment itself there were many inaccuracies, also regarding the scientific literature. Divergence was diametric, since the basic assumptions are taken into consideration and some articles presented completely different point of view than others. It had to be checked and that took a lot of time.

The second difficulty regarded the proportion of nanotubes. Shall it be 0.5% or rather 5%? In the early experiments, in which I started to combine nanotubes with the epoxy resin, agglomerates started to form. They're structures visually similar to coffee grounds. We had to stick to one direction and analyse it from the smallest to very high values, to find extremes in an experimental way. Another challenge: the manufacturer declared that nanotubes had a diameter of 6 to 9 nanometres. But the first measurements using a scanning microscope have given the results of 69 to 70 nanometres, so the inaccuracy reached an order of magnitude. Only after several months of simulations and trying to find the errors or matters not taken into consideration, I took photographs on a transmission microscope. It proved that the declared nanotubes' size was correct, whereas the scanning microscope hadn't been precisely calibrated. I had to continue research using other equipment.

By defending your thesis you have closed some stage. Could you share some plans for the future?

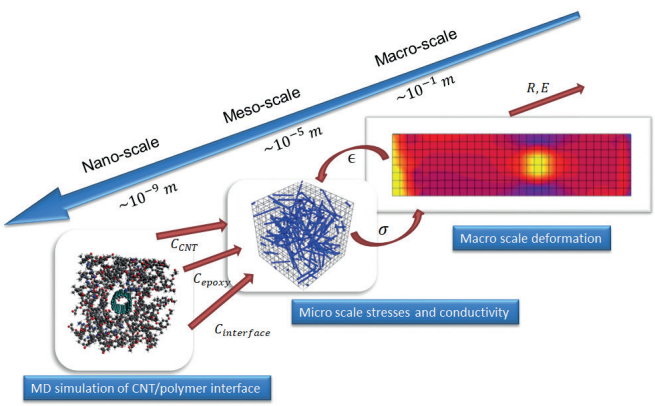
When I was about to finish work on my PhD project, I told my supervisor, that I felt like I knew much less than I had known at the beginning. After several years I became much more aware of the quantity and variety of possible paths. Since then I have made several serious attempts to create intelligent sensory structures. For example, together with my colleague we created methane-detecting sensors. In a mine we can cover the spans with such a structure and not only we will be alarmed about a strain, but also we will be able to measure concentration of methane, which is very dangerous for miners.

You have mentioned, that after finishing PhD studies you had recognised many new paths. What could you advice to younger colleagues who are at the beginning of their scientific career?

I would primarily suggest to focus on one single thing and don't be worried about failures. Because, as for me, working on the good PhD thesis is littered with failures. That's the scientist's work. It's a kind of risk. The thesis does not have to necessarily finish successfully, instead it should show the way.

I would also like to advise young scholars not to discourage if they find some similar works. In my opinion, even if you notice, that someone's research is going in the similar way, it is a good idea to write an e-mail, provide comments to the other author, give some feedback to help somewhat. Personally, I was afraid of acting this way at the beginning, but now I don't hesitate to do that. Very interesting contacts may be made as a result.

The science is a startup. We receive grants, and investing in grants is like investing in startups. Some of financed projects will bring effects, some won't. Nevertheless, the process itself, the way, may help to discover something completely new.



Multiscale model of CNT/polymer sensors (coatings) – graphical representation

Agnieszka Kącka-Zych

The interview with the author of the PhD thesis:
„Theoretical study on the mechanism of carboxylic acids elimination from esters of nitroalcohols".

How long have you been interested in chemistry (in broad meaning) and how did this interest begin?

I got interested in chemistry when I was in secondary school. In the first year at university I became very excited in chemistry, especially organic chemistry. During my studies I worked in a Scientific Circle of Chemists led by Professor Radomir Jasiński, as it turned out, my future supervisor of the doctoral thesis.

What did you expect from your PhD studies at their beginning? Did these expectations come true?

Before the PhD studies I expected first of all personal development, broadening the horizons and that the doctorate would allow me to perceive new phenomena and existing ones to perceive differently. These expectations have come true.

What influenced your decision to do interdisciplinary research on the border of chemistry, physics and mathematics?

Interdisciplinary research is where two or more disciplines combine to create something more than the sum of their parts. It is about not being constrained by one way of thinking or tackling a problem. Instead, skills and knowledge are gained in a variety of disciplines and new applications for the research are sought.

Could you point some practical implementations of conjugated nitroalkenes that are the subject of your study?

Conjugated nitroalkenes have proved to be a valuable group of reactants in organic chemistry. Their strong electrophilic character makes them important precursors to a wide variety of target molecules. They are useful compounds applied in organic synthesis. The applications of nitroalkenes in organic synthesis is largely due to their ease conversion into a variety of functionalized derivatives. Moreover, these compounds were found to be important because of their biological activities. Among various biological properties, the anticancer activity of nitroalkenes and their novel MBH adducts with other activated alkenes has highlighted the enormous potential of nitroalkenes' derivatives as bioactive molecules. Furthermore, nitroalkenes are important precursors of many insecticides, fungicides and pharmaceuticals.

Some of the methods you have used during your research were pioneer in Poland. Could you introduce them briefly?

These are the methods proposed by Professor Luis Domingo, with whom I had the opportunity to

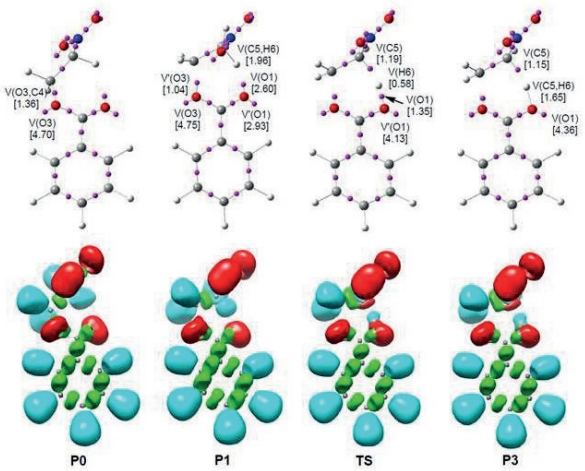
work as part of the preparation of my doctoral dissertation at the University of Valencia. Molecular Electron Density Theory (MEDT) is based on the idea that while the electron density distribution at the ground state is responsible for physical and chemical molecular properties, as proposed Density Functional Theory, the capability for changes in electron density, and not molecular orbital interactions, is responsible for molecular reactivity. Therefore, the reactivity in organic chemistry cannot be characterized neither by a static energy nor by a geometrical study of the corresponding stationary points, including the transition state structures, but by a rigorous analysis of the molecular electron density changes along the reaction path, as well as of the changes in energies required to reach the transition state geometry.

Your research required much calculation and data-processing. In what way were ACC Cyfronet AGH resources helpful with them?

Actually, all my doctoral dissertation is based on calculations made on the Prometheus supercomputer located at the ACC Cyfronet AGH. As I have already mentioned, I also conducted theoretical research using the Molecular Electron Density Theory, which combines Electron Density Theory and Bonding Evolution Theory. These calculations would not have been possible without the help and commitment of the employees of Cyfronet, for which I would like to thank them very much.

Could you share some advice with young researchers who are starting their PhD studies?

Don't give up on the pursuit of the goal, and treat any errors or tripping as valuable lessons. Take advantage of international travel offers, student exchanges or trips to conferences. Get to know as many inspiring people as possible and always try to learn something from them. Be open to the world and the opportunities that stand before you.



The stages of decomposition of nitroethyl benzoate, in the light of the Electron Localization Function (ELF) theory

Kamil Kurleto

The interview with the author of the PhD thesis:
„Theoretical studies of structure and olefin metathesis activity of molybdenum and tungsten oxide species on silica".

How would you explain your research to a non-specialist?

I have some favourite comparison explaining why we have made these computations. Everyone involved in chemistry experiments knows, what he puts inside the flask, and knows or rather is able to predict what he will take out from it. But we are pretty unsure, what happens inside this reaction. The calculations I perform allow us to analyse the reaction mechanism, consequently understand the process better, which may lead to improving it. It's worth noting that calculations aren't made separately from the real processes. They're strictly connected to the experiment. Of course, theoretical and experimental research can be made individually, but only after combining them we become able to obtain the complete information.

Although olefin metathesis, which is the subject of your research, is also an interest of big companies, you have found a field, in which so far there were more questions than answers.

As for generating olefins, metathesis process, I mean a carbene mechanism, is recognized quite well. But so far it hasn't been explained, how the active centres of catalysts used in industry are formed. The whole process can be divided into three stages: initiation, propagation and deactivation. We already know, how the propagation stage goes, but it hasn't been explained yet, what mechanism lays under initiation (forming of active centres) and deactivation. That's why I decided to focus on these two stages in my research.

Could you point some implementations of olefin metathesis?

Olefin metathesis reaction is, above all, used in industrial processes of propylene production. We can also find other implementations in oleochemistry, agrochemistry, pharmacology and aromatic industry.

What were the biggest challenges during your research?

I think that the biggest challenge was to create an appropriate, realistic enough catalyst model. It was so important, because the more precise results of calculations we get, the easier it is to compare them to results of the experiment. Another challenge was to analyse a vast number of virtual paths of olefin metathesis initiation stage, not only described so far, but also proposed by myself.

In which way were the ACC Cyfronet AGH resources helpful in creating a new functional model?

Cyfronet has helped me primarily by providing access to huge computing power, guaranteed by Zeus and Prometheus supercomputers. One of the advantages of computational chemistry is that

the calculations can be performed almost in every place equipped with a computer and required software. However, we lose a lot of precious resource, which is time needed to do that. Thanks to Cyfronet's supercomputers I was able to perform calculations easily and fast, what brought my PhD thesis as a result.

What further steps can be made on the basis of your research?

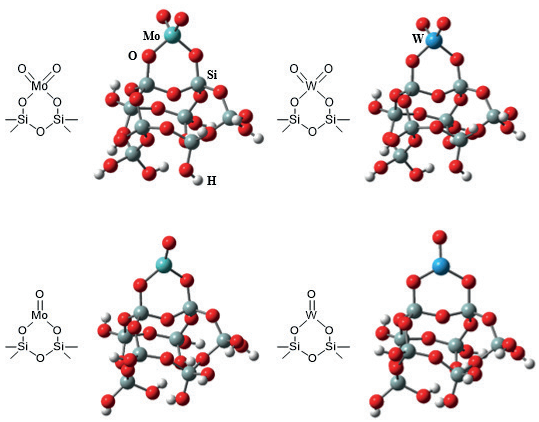
The grant we received included two parts. Our sub-team performed calculations whereas at the same time the other sub-team made experiments. Our task was to propose new olefin metathesis catalysts models, to analyse the mechanism of this process, and to perform theoretical calculations simultaneously to experimental research. Eventually, creating a new, more efficient catalyst system for olefin metathesis process may be an outcome of this effort. Due to the fact that in my thesis I focused on analysing molibdenum and tungsten catalysts on silica, I think that the natural consequence should be to perform calculations for these catalysts on aluminosilicate or aluminum oxide supports.

What aspect of your work is most satisfying?

If a challenging path of synthesis or calculations causes a lot of problems, the satisfying final result always gives me a lot of joy. I also appreciate the possibility to develop my skills constantly. In my MSc thesis I analysed antioxidant activity of chosen plants, during PhD studies I obtained knowledge and skills regarding olefin metathesis and computational chemistry. And actually, I work as a synthesis specialist. This variety of research subjects I'm involved in, is very important to me and gives me a lot of satisfaction.

By defending your PhD thesis, you have closed some stage. What advice could you give to people, who have already started it?

I would advise the people, who haven't started PhD studies yet, to choose the supervisor wisely. Furthermore, If I were to share advice with students already involved in PhD research, I would suggest to focus on your thesis and put an effort to work on it. At the same time, if a possibility to start cooperation, to take a scientific internship or to work in industry arises, it's something you should benefit from. I would also encourage everybody to take an active part in scientific conferences. It always gives the possibility to present the outcomes of your work, become more familiar with the people working in related fields and to improve foreign language skills.



Structures of the studied in the dissertation catalytic systems for olefin metathesis process (oxide precursors Mo(IV), Mo(VI), W(IV) and W(VI) deposited on the surface of amorphous silica)

Alina Mreńca-Kolasińska

The interview with the author of the PhD thesis: „Simulations of quantum transport in graphene systems with n-p junctions".

In recent years graphene has become very popular in the world of science. What influenced you to devote your dissertation to this material?

I devoted my whole PhD project to graphene, because at the time it became a very up-to-date and rapidly developing subject, due to its unusual properties. I got interested in graphene mostly because the Nobel Prize had been awarded for its isolation – especially that the scientists had done it in the very simple way, by sticking and tearing off a tape. Among removed layers, many had thickness of one atom, and it was graphene. Since my supervisor wanted to start research on graphene at the time, it was a great chance for me to get to know more about this material's properties.

In which way may the analysis of phenomena occurring in graphene in the presence of scanning-gate be implemented in a practical way?

As far as I know the method is currently used for basic examination of graphene's properties. Nevertheless, I suppose that one of the phenomena observed in n-p junctions in graphene, namely fractional quantization of resistance, could be used as e.g. precise standard of electrical resistance. Such junctions can be created inter alia with the use of scanning-gate.

What further challenges stand in the way of scientists investigating graphene?

Electronic and mechanical properties of graphene are very promising. But the material also has its weakness, which is very weak spin-orbit coupling. It can be used to control spin dynamics, thanks to which it would be possible to use spin as an information carrier – similarly to particle charge, which is such a carrier in electronics. It's the object of interest of spintronic. At the moment there's a research going on, aiming to increase the spin-orbit interaction in graphene and I see it as one of the greatest challenges in this field.

In which way did Cyfronet's resources help in performing calculations and in the analysis of their results that was necessary in your research?

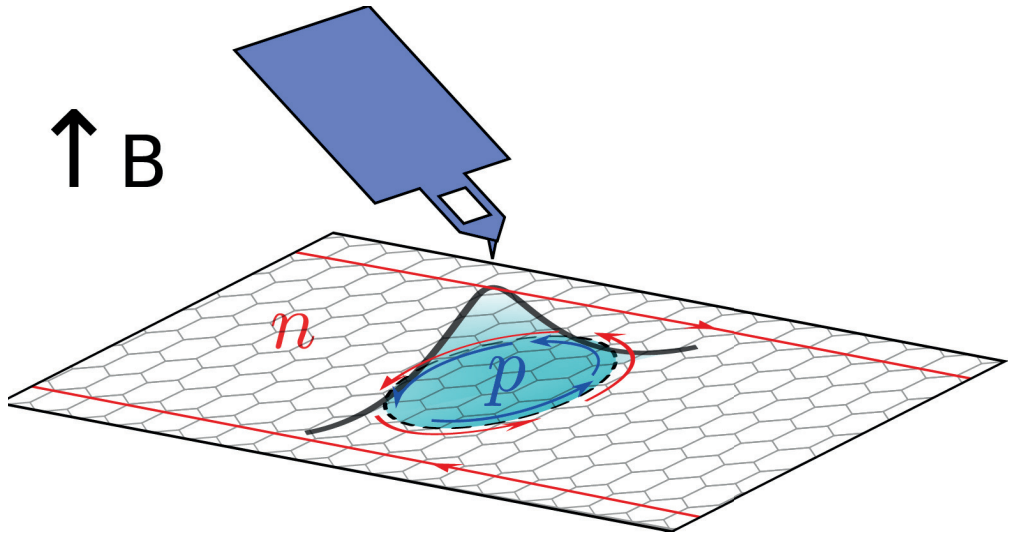
They were invaluable in conducting my research. To be honest, without them it would be impossible to get the most interesting results, especially those related to simulation of scanning-gate technique. These calculations required so much time, that without access to Cyfronet's infrastructure, I would probably have been making simulations on faculty's computer until now.

What aspect of your work gives you the greatest satisfaction?

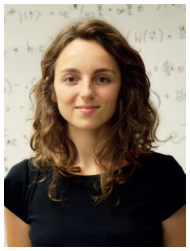
I find the moments of hammering out the solutions of problems and enigmas the most joyful. For example, during the intern in the experimental group, the result of a sample measurements we got at first seemed to be completely beyond reason or even mistaken. After spending much time on thinking it through, many disputes and literally hours of calculations with a pen and a piece of paper, we finally managed to create a model of electrons' and holes' behaviour in graphene. Its result was fully compatible with our measurements. It was the moment of the greatest satisfaction I have experienced during my PhD studies.

What could you advise people, who are at the beginning of PhD studies? What should they focus on?

I would like to advise PhD students to try to learn something new all the time, despite the research itself is much time-consuming. It's good to broaden your horizons, especially in the fields related to the one you investigate. This knowledge may surprisingly become extremely useful for interpreting our own results. Moreover, it's good to try to be up-to-date with the recent work of other scientists, e.g. follow new publications and the development of our field of research.



A circular n-p junction induced within a graphene nanoribbon by an external potential of a scanning probe



Edyta Osika

The interview with the author of the PhD thesis: „Dynamics of spin-valley transitions in carbon nanotube quantum dots".

How did your interest in nano-physics begin?

It all started with my engineer thesis – I was looking for a project in theoretical physics in which I could also develop my programming skills. That is how I began working with Prof. Bartłomiej Szafran on simulations of the processes taking place in quantum dots. Very soon I got so interested in the topic that I continued working on it during my master and doctoral studies.

What influenced you to devote your PhD thesis to an analysis of electrical processes that may be identified in carbon nanotubes?

A few months before I started my PhD studies, a scientific group from the Netherlands carried out an experiment performing successful spin rotations within carbon nanotube quantum dots. At that time we have already had some experience with analysis of similar processes within semiconductor nanowires. The results of the experiment, however, could not have been explained using the already existent theory. As a graphene-based material, carbon nanotubes provide much more complex environment for single spins manipulation than the semiconductors we investigated before. On the other hand, thanks to the absence of a hyperfine field, carbon nanotubes are truly unique in terms of their potential of using in future quantum computers.

Although quantum information science may seem to be a futuristic discipline, it is being developed and discussed in many research centres. What role does your research play in this international debate?

Our research is focused on devices, which could be used to build a qubit – a quantum analogue of a classical bit. Qubit is a very fundamental part of a quantum computer, but at the same time the greatest problem for quantum information science – the decoherence – is directly related to it. Great effort is being made nowadays to find materials, which can provide the longest coherence time – that is the time the qubit can store some information without destroying it. Carbon nanotubes are really promising material in that aspect, that is why we are doing our best to widen our knowledge about them.

How close are we – in your opinion – to the moment, when quantum computers will not only be functional, but also easily accessible? What still needs to be done?

We are at the very beginning of making quantum computers accessible for everybody. There are plenty of issues, which have to be approached, starting from technological challenges – control over decoherence, efficient qubits entanglement – finishing on software development – creating a language which would allow to use quantum algorithms without deep knowledge of their theoretical fundamentals. I do not think we will be able to buy a quantum computer to use it at home anytime soon – not in next 10-20 years. I believe, however, that we will get access to quantum computers

developed and built in the scientific centres in the world. I can imagine we will be able to use them for scientific research at similar terms as we nowadays use supercomputers, e.g. the ones at the ACC Cyfronet AGH.

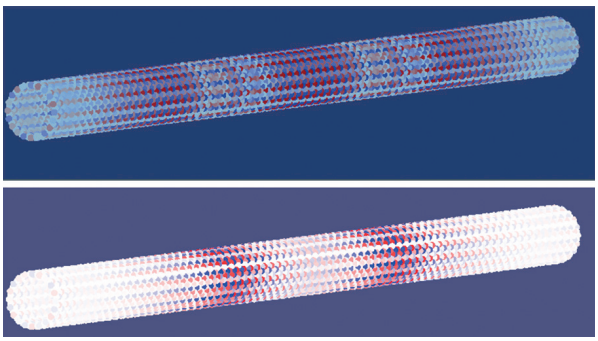
Which aspect of your research was the most challenging? What were the biggest obstacles to overcome, and – on the other hand – what did you find most satisfying?

Definitely, the greatest challenge in our research was dealing with insufficient computational power to carry out all the simulations we would like to perform. The majority of our simulations were so time- and computing power consuming, it would be impossible to run them using only personal computers. In that context, the possibility to use Cyfronet infrastructure has been really invaluable help for us. But still, the computing power limitations were the most important factor determining how far we can proceed with our research.

The most satisfying were the moments when we could present our research to a wider audience, e.g. at conferences or through publications in scientific journals. Seeing that our research is useful and interesting for other scientists is the best reward for the months of hard work on collecting and analysing our results.

What did you use the ACC Cyfronet AGH infrastructure for?

Our main goal was to create a model of quantum dots within carbon nanotubes and use it to simulate processes driven by oscillating electric field. We used the atomistic approach to achieve it. This kind of approach allows to skip several approximations – which are usually made in similar calculations – and obtain better, more exact results. But this comes with a price – a huge computational cost of calculations. The ACC Cyfronet AGH infrastructure was irreplaceable here – it shifted our limits to the completely new place. Simulations, which thanks to the parallelization of the processes took us few weeks on supercomputers, would not be even possible if working only on personal computers.



Wave function of an electron delocalized (upper picture) and localized within carbon nanotube double quantum dot (lower picture)

Basing on your experience as a scientist, what advice could you share with researchers at the beginning of their career?

In my opinion, it is very important to – starting from the very beginning of our carrier – always choose topics, which we are truly interested in. Only the curiosity and authentic engagement in the research can turn our day-to-day work into a truly fascinating adventure. Then, it is easy to stay motivated to do your best at work, it is easy to have the patience and perseverance necessary to succeed. And even if the success does not arrive as fast as we would like to, we can always find satisfaction just in the process of pursuing it.

Kacper Pilarczyk

The interview with the author of the PhD thesis: „Information processing in molecular-scale systems based on carbon nanostructures".

What influenced your decision to become a scientist?

I have been interested in chemistry and physics since high school and my first contact with some actual scientific problems was through the workshops organized by the Polish Children's Fund. I remember that the decision to become a scientist came after the internship during which I have been supervised by prof. Marek Pietraszkiewicz, whose attitude, expertise and valuable advices confirmed me that being a researcher, although demanding, can be really satisfying and enjoyable.

How would you explain the matters you are working on to a non-scientist?

We rely greatly on devices capable of storing and processing information. It is natural that we expect a constant increase in the performance of these. This progress is possible inter alia due to miniaturisation. At the same time, we know that there are some limits along this path. For that reason, there is a need to search for alternative approaches which may be based on the diversification of used materials or the change of paradigms based on which our devices work.

Your research combines elements of chemistry, nanotechnology, photo- and biophysics, logic and some other disciplines. Why is it so important to cross the borders of various disciplines?

We should avoid rising artificial borders and search for the similarities between different fields instead. If we try not to think about our disciplines as excluded beings, but rather agree that they are a part of a broader picture, we will notice how many ideas from other areas can be adapted in our own research. This is especially important if we want to answer some of the fundamental questions in science and address some of the practical problems – we can see it in medicine, where new therapy methods are based on the use of molecules specially designed by biologists and chemists.

Could you point some milestones of your research?

One of the most important milestones was when I got involved in the research on the information processing realised in chemical systems and the unconventional computing, which happened around the second year of my PhD course in the group of prof. Szaciłowski. At that point I got interested in this field and the first article on the ternary logic implemented with the use of nanocomposites appeared in 2015. We discussed the possibility of using an approach different than the classic binary logic in a rather simple optoelectronic device and that was pretty exciting. We went further in 2016, when we presented a simple light-driven system that was capable of mimicking basic elements constituting a nervous system – synapses. The last milestone I can think of was a book chapter and a review paper in which we suggested that our materials could be used to implement fuzzy logic systems.

How close are we, in your opinion, to construct artificial neural circuits?

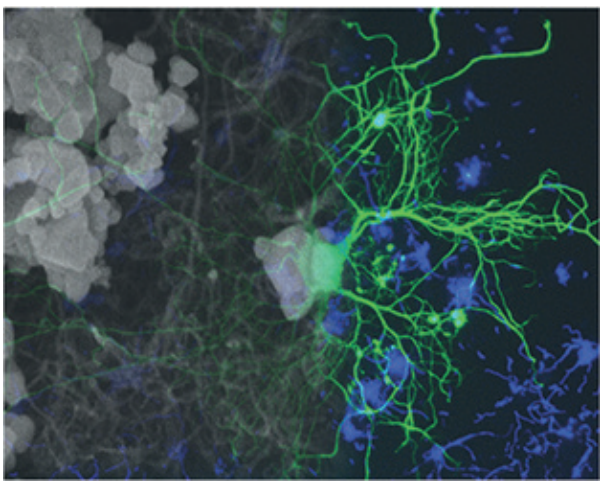
It is happening right now and I mean not only computer simulations – which are even more developed at the moment – but also hardware realisations of such systems. Nowadays, memristors-based solutions are quite popular and researchers try to achieve information of processing patterns similar to that characteristic for biological structures. Nevertheless, these are usually simple networks and I am pretty sure it will take some time to get even close to the complexity of a human brain, not to mention its brilliant cognitive abilities. Still, we may use these minimalistic models to better understand mechanisms behind the information processing and storage realised by our nervous system or use these circuits to perform some basic automated actions with a taste of adaptability characteristic of our own behaviour.

Why did you decide to use multi-valued and fuzzy logic systems instead of binary one?

As for the ternary logic we believed it is a more natural way of analysing states of the systems based on the interaction with light. I have mentioned on several occasions that it is pretty straightforward to encode three different states when we use light pulses – one is associated with the 'off state' and another two with two different polarisations. And the interest in the fuzzy logic resulted probably from our works on neuromimetic systems. This concept is strongly related to the way we process pieces of information so we wanted to check if it is reasonable to apply it in the case of our devices, at first with a rather simple example and then, in the paper published last year, for a more complex case.

What role did the ACC Cyfronet AGH resources play? How did they help with working on your thesis?

In order to test our ideas and to design optoelectronic devices implementing the multi-valued logic or some aspects of neuromorphic engineering, we also work on new hybrid materials, nanocomposites, which physicochemical properties are tuned in an appropriate manner. To better understand phenomena responsible for materials' properties or to predict what we have to change in the system to achieve a certain behaviour, the DFT modelling and other numerical methods are of a great assistance. Obviously, the more complex the system is, the more resource-demanding are the computations, thus we need Cyfronet's supercomputers to obtain the results within our lifetime.



Simulation of attempt to reconstruct nerve connections (on the right) with the use of nanocomposite (on the left)

Basing on your experience, what advise could you give to PhD candidates and students?

To be passionate about what they do irrespectively of the discipline they choose and to have fun with it. That is very important for the creativity. And to be open-minded – the inspiration can come from any direction.

Supercomputers for AGH Racing

Bartosz Wróbel
AGH Racing

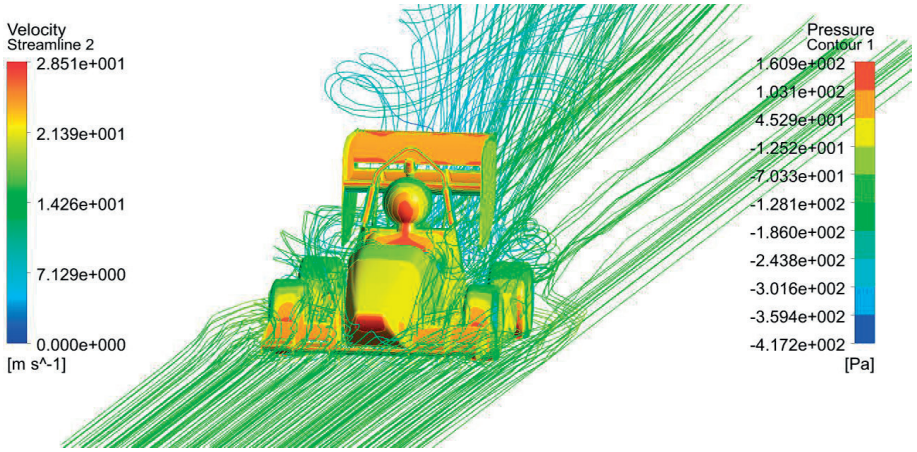
AGH Racing is a project carried out by the group of students consociated in the academic circle called Kinematics, running on the Faculty of Mechanical Engineering and Robotics. The main aim is to design and manufacture a racing car for international competition Formula Student. To meet the challenge, the team is divided into sub teams that deal with: aerodynamics, chassis, electronics, powertrain, suspension and marketing.

One of the most demanding task is numerical analysis of airflow which is crucial for designing an aerodynamic package that consists of front, rear and side wings. These elements are responsible for generating additional downforce that improves traction what results in shorter lap times during racing.

Design of aerodynamic package was possible due to professional ANSYS Fluent software and supercomputer Prometheus available at ACC Cyfronet AGH. Thanks to quick and easy access in interactive mode through TurboVNC client, the geometry of the racing car and computational domain was developed. Consecutively it was verified in grid independence test thereupon the computational power was used effectively.

The design phase started from creating parametric model and running airfoils setup analysis after their choice and adaptation. The next step was founding their proper position in the racing car in reference to rules requirement. Interaction between each of them was also examined to gain smooth distribution of streamlines. The wing's shape was gradually modified to reduce local eddies of the air towards minimizing drag. Optimisation also allowed to compensate pressure gradient resulted from asymmetry of the car. The key role in design phase was to determine assumed distribution of aerodynamic forces acting on front and rear axle. Setup of the wings can significantly appeal to car's performance that's why the special, low drag configuration was prepared for acceleration event. Finally, to assess mileage of the aero package, the comparison to a non-wing car was conducted.

Summing up, taking advantage of high performance computers let me complete design for aerodynamic package. Ability to perform calculation in parallel sessions allowed to consider a wide array of different configurations in short period of time. On the basis of the best one the change of aerodynamic forces against speed of the race car was plotted.



Distribution of pressure and airflow around the car model at 15 m/s

AGH Racing Car at TT Circuit Assen racetrack, The Netherlands

CFD in AGH Solar Boat

Marcin Pająk, Monika Wolny
AGH Solar Boat Team

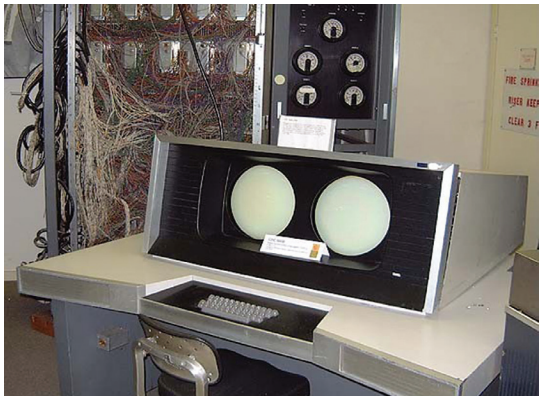
AGH Solar Boat team is a group of people contributing to the development of the Solarboat vessel, which takes part in international championships called Solar Sport One and Monaco Solar & Energy Boat Challenge. The whole idea is to design and build a watercraft, which is powered from photovoltaic cells. Members of the project come from many different students' organizations existing alongside the AGH University of Science and Technology, but the whole idea originated from the "Eko-Energia" scientific circle from the Faculty of Energy and Fuels as well as the AGH Academic Sailing Club. Students of the Faculty of Mechanical Engineering and Robotics also make an enormous contribution to the project, and more and more of them have joined it lately. The workforce is distributed in smaller teams, each working on a certain field of the vessel development, such as construction, electronics or marketing.

However, the team responsible for construction was the one who used the ACC Cyfronet AGH resources at most. The CFD simulations are very vital for the project, as they help to identify if the current concept of the boat's design is heading in the right direction or if certain corrections are required. The analyses conducted within the project concerned hydrodynamics, heat transfer and durability. Due to their high complexity, those analyses required tremendous computing power and it was impossible to perform them on a single PC. This fact determined using high-end supercomputers, which was possible by courtesy of ACC Cyfronet AGH. The software used by the team members was Ansys 18.2, available at the Prometheus supercomputer.

The exact subject of analyses conducted using the shared resources was a set of hydrodynamics-focused simulations of flows over the nacelle and hydrofoils. The complexity of designed geometries required application of a very dense mesh. That fact determined long computation time and necessity of usage of parallel computation, running on all 24 cores available in a single Prometheus node. Obtained results served for finding parameters of our biggest interest, which were drag and lift coefficients. The latter one especially matters in case of hydrofoils. The movement of our solar vessel is based on their application. Their task is to rise the boat's hull completely out of the water, while only the foils are submerged. This procedure allows a significant decrease in friction produced during the vessel's motion. However, this effect can be achieved only if the lift coefficient is high enough. Moreover, the generated lift is directly correlated with the boat speed. So, the analyses also helped us find the minimal velocity required to rise the vessel out of the water and, at the same time, sufficient enough to maintain this state. As long as a solar-powered watercraft is considered, every operation leading to increase of the energy efficiency, like mentioned decrease of friction, is very desirable and expected. The more energy we save, the longer distance can be covered by the watercraft.

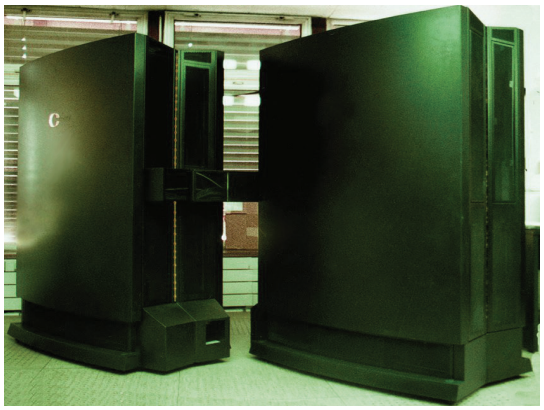
AGH Solar Boat “Baśka”

AGH Solar Boat team during Monaco Solar & Energy Boat Challenge



CDC CYBER 72

Convex C3840



Exemplar SPP1600/XA

1973	CYFRONET is established.
1975	A CDC CYBER 72 computer is deployed at the Centre.
1990	The first KraKow node of the EARN / BITNET network is deployed at CYFRONET (on an IBM 4381 computer).
1991	CYFRONET installs a Convex 120 machine – the first vector computer in Central and Eastern Europe. The first Polish national Internet link is established between Krakow and Warsaw. Construction begins in the Krakow MAN.
1994	A 2 Mbps link is deployed between Krakow and Warsaw.
1996	An Exemplar SPP1600/XA computer deployed at CYFRONET took a position on TOP500 list. The first automatic tape library (ATL 2640) is installed at the Centre.
1997	The ATM communications subnet is deployed within the Krakow MAN. CYFRONET joins the POL-34 national backbone.
1998	An SGI Origin2000 computer is deployed at the Centre.
2002	A RackSaver PC cluster is deployed at CYFRONET as part of the CrossGrid project.
2003	An HP Integrity SuperDome computer is deployed at CYFRONET (the first such computer in Poland).
2005	An HP Storage Works XP12000 disk array is deployed at CYFRONET.

2006 An HP Storage Works EVA 8000 disk array and an SGI ALTIX 3700 computer is deployed at CYFRONET.

2007 An agreement concerning the creation of the Polish Grid (PLGrid) Consortium was signed.

An SGI ALTIX 4700 computer is deployed at CYFRONET.
Two IBM BladeCenter HS21 servers are deployed at CYFRONET.

An HP Storage Works EVA 8100 disk array is deployed at CYFRONET.

SGI Origin2000

2008 The configuration of SGI ALTIX 3700 computer is extended to 256 processors.

MAN 10 Gbps started.

Zeus supercomputer (HP Cluster Platform 3000 BL) is deployed at CYFRONET.

SGI ALTIX 3700

2009 Start of the PL-Grid project – Polish Infrastructure of Supporting Computational Science in the European Research Space.

2010 The extention of the configuration of HP Cluster Platform 3000 BL to 9544 cores (Intel Xeon processors) allows CYFRONET's supercomputer to rank 161st position on the TOP500 list.

2011 Deployment of Hitachi Data Systems High Performance NAS for computing infrastructure.

Total amount of installed disk space exceeds 2 PB.

The configuration of Zeus supercomputer (HP Cluster Platform 3000 BL) is extended to 12,032 cores (Intel Xeon processors).

Zeus supercomputer has been placed on 81st position on the TOP500 list.

HP Cluster Platform 3000 BL

Anniversary Medal

New Machine Hall

Prometheus supercomputer

2012 Start of the PLGrid Plus project – domain-oriented services and resources in the PL-Grid.

In April, ScaleMP, a leading provider of virtualization solutions for high-end computing, announced that Zeus-vSMP system at CYFRONET is the largest virtual SMP system in Europe.

Zeus among 100 fastest computers on the TOP500 list.

2013 After upgrading of Zeus supercomputer configuration, the theoretical peak performance reached 374 TFlops.

Anniversary Medal has been minted.

2014 The new Machine Hall is completed.

Start of two new projects: PLGrid NG and PLGrid Core.

2015 The Prometheus supercomputer is deployed at CYFRONET, and ranks high, 49th place in the July edition, and then 38th place in the November edition of TOP500 list.

For the first time in history two supercomputers from CYFRONET (Prometheus and Zeus) are ranked on the TOP500 list, in one edition.

The new backup Data Center is completed.

CYFRONET starts active participation in INDIGO-DataCloud, EGI-Engage, EPOS- IP and PRACE 4IP projects.

2016 Prometheus ranks 48th (the June edition) and 59th place (the November edition) on the TOP500 list.

2017 Prometheus ranks 71st (the June edition) and 77th place (the November edition) on the TOP500 list.

Further dynamic development, including establishment of 6 new laboratories.

Sat4Envi, Gliomed, EPOS-PL and eXtreme DataCloud projects launched.

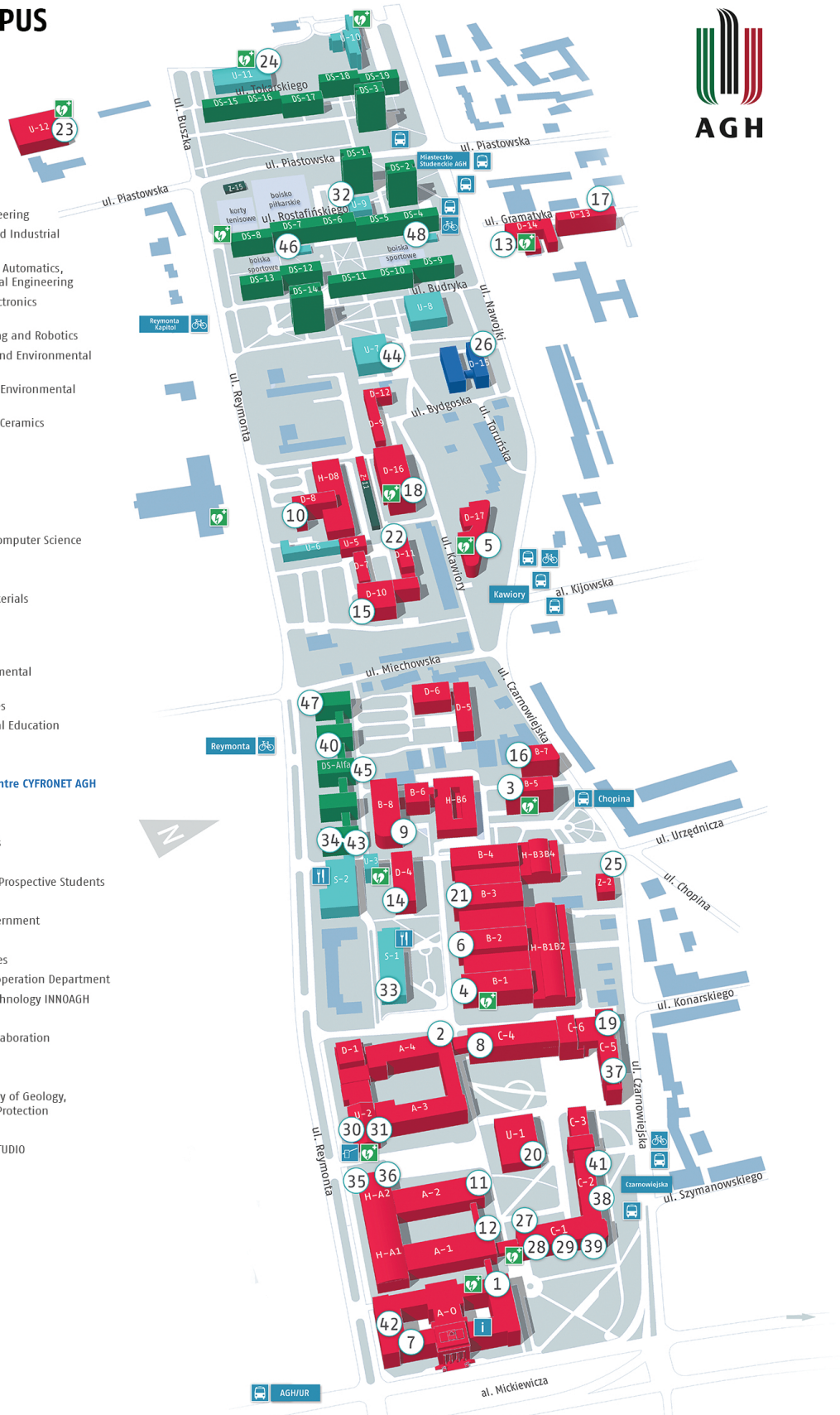
2018 Prometheus ranks 103rd place on the TOP500 list (the June edition).

EOSC-Hub project has been launched.

AGH UST CAMPUS

LEGEND

1. Rector's Office
2. Faculty of Mining and Geoengineering
3. Faculty of Metals Engineering and Industrial Computer Science
4. Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering
5. Faculty of Computer Science, Electronics and Telecommunications
6. Faculty of Mechanical Engineering and Robotics
7. Faculty of Geology, Geophysics and Environmental Protection
8. Faculty of Mining Surveying and Environmental Engineering
9. Faculty of Materials Science and Ceramics
10. Faculty of Foundry Engineering
11. Faculty of Non-Ferrous Metals
12. Faculty of Drilling, Oil and Gas
13. Faculty of Management
14. Faculty of Energy and Fuels
15. Faculty of Physics and Applied Computer Science
16. Faculty of Applied Mathematics
17. Faculty of Humanities
18. AGH UST Academic Centre for Materials and Nanotechnology
19. AGH UST Centre of Energetics
20. Main Library
21. Walery Goetel School of Environmental Protection and Engineering
22. Department of Foreign Languages
23. Department of Sport and Physical Education
24. AGH UST Swimming Pool
25. Centre of e-Learning
26. **AGH UST Academic Computer Centre CYFRONET AGH**
27. University Computer Centre
28. Department of Education
29. Centre for International Students
30. Education Centre
31. University Admissions Board for Prospective Students
32. AGH UST Student Campus
33. University Board of Student Government
34. Career Centre
35. Centre for Transfer of Technologies
36. Administration and Business Cooperation Department
37. Krakow Centre for Innovative Technology INNOAGH
38. Centre for Project Management
39. Department of International Collaboration
40. Disability Support Office
41. AGH UST Museum
42. Geological Museum of the Faculty of Geology, Geophysics and Environmental Protection
43. AGH UST Press
44. Academic Cultural Centre, Club STUDIO
45. Student Club Gwarek
46. Student Club Zaścianek
47. Student Club Karlik
48. Student Club Filutek



Academic Computer Centre
CYFRONET AGH
Nawojki 11
30-950 Krakow.P.O.Box 386

phone: +48 12 632 33 55
fax: +48 12 633 80 54

cyfronet@cyfronet.pl
www.cyfronet.pl

ISBN 978-83-61433-29-3